IDENTIFICATION AND CHARACTERIZATION OF A DNA 5^{m}C-CYTOSINE METHYLTRANSFERASE GENE IN STREPTOCOCCAL CONJUGATIVE TRANSPOSON, Tn*5252*

By

JANARDHAN SAMPATH

Bachelor of Pharmacy
Birla Institute of Technology and Science
Pilani, India
1990

Master of Science
Birla Institute of Technology and Science
Pilani, India
1990

Submitted to the Faculty of the
Graduate College of the
Oklahoma State University
In partial fulfillment of the
Requirements for
The Degree of
DOCTOR OF PHIOLOSOPHY
December, 1998

IDENTIFICATION AND CHARACTERIZATION OF A DNA

5^{m}C-CYTOSINE METHYLTRANSFERASE GENE

IN STREPTOCOCCAL CONJUGATIVE

TRANSPOSON, Tn*5252*

Thesis Approved:

M. N. Vijayakumar

Thesis Advisor

Robert D. Hotchkiss

Ulrich Melcher

Melanie J. Palmer

Jeff J. Washington

Wayne B. Powell

Dean of Graduate College

DEDICATION

To my mom for her great sacrifices

to make me who I am.

ACKNOWLEDGEMENTS

My keen appreciation goes to my major advisor Dr. Moses N. Vijayakumar for teaching me science, for his patience, intelligent guidance, encouragement, enthusiasm and a true friendship. I would also like to thank my committee members Dr. Jeffery Hadweiger, Dr. Ulrich Melcher, Dr. Robert V. Miller and Dr. Melanie Palmer for their constructive advice during my research work.

I am also grateful to my former colleagues, Dr. Ali O. Kilic and Dr. Sahlu Ayalew for their keen interests during the initial stages of my work and giving the right directions to proceed.

My sincere thanks extend to my former colleague Dr. Francisco J. Alarcon and my colleagues Ursula Munoz Najar, Kanchana Natarajan, and Jaishree Srinivasan for the valuable discussions, help, and friendship.

I wish to thank Dr. Kay Scheets for providing me with antibodies and for interest in my scientific problems.

A special thanks is extended to Charlotte Regson, Elfriede Miller, Tara Bennett, Audrey Sheridan, and Sandra Peterson for their help.

I would also like to thank the staff and members of the Department of Microbiology and Molecular Genetics for their support and for providing a friendly environment.

In particular I wish to thank my wife, Anjali for her understanding, patience, and for great company and friendship.

I dedicate this work to my mom who has been a continuous source of inspiration throughout my life and whose belief in me has brought me to this stage.

TABLE OF CONTENTS

Chapter Page

I. INTRODUCTION 1

Regulation of gene expression 3

Tn*10* 3

Tn*5*/Tn*903* 4

Bacteriophage Mu *mom* gene 5

Bacteriophage P1 *cre* gene 5

Bacteriophage P1 *pac* cleavage 6

Anti-Restriction 7

Specific aims 11

II. LITERATURE REVIEW 12

The Organism: *Streptococcus pneumoniae* 12

Drug Resistance in *S. pneumoniae* 13

Source of Antibiotic Resistance 14

Cm^r and Tc^r are Chromosomal Insertions 15

Ω*cat tet* is Conjugative Transposon 16

Uniqueness of Conjugative Transposon 17

Chapter Page

Classification of Conjugative Transposon....17

Conjugative Transposons of Gram Negative Bacteria....18

Conjugative Transposons of Gram Positive Bacteria....19

Tn*916*/Tn*1545* family....19

Tn*916*....20

Tn*1545*....21

Tn*5253* family....22

Tn*5276*....22

Tn*3701*....23

Tn*3951*....23

Tn*5253*....24

Tn*5252*....27

DNA Methyltransferases....29

Type-I....29

Type-III....31

Type-II....32

III. MATERIALS AND METHODS....38

Bacterial Strains....38

Growth and Storage of Bacterial Strains....38

Plasmids....43

Chapter Page

Enzymes, Chemicals and Antibodies....46

Transformation....50

S. pneumoniae Transformation....50

E. coli Transformation....51

Conjugation....52

Chromosomal and Plasmid DNA Isolation....53

Chromosomal DNA isolation from *S. pneumoniae* 53

Plasmid DNA isolation from *S. pneumoniae* 54

Plasmid DNA isolation from *E. coli*....56

DNA Manipulation....57

Cloning....57

Agarose Gel Electrophoresis....58

Conversion of 3' Overhang to Blunt-End with Klenow Fragment 58

DNA Sequencing with T7 DNA Polymerase 59

Denaturing Polyacrylamide Sequencing Gel....59

Construction of Nested Deletions for Sequencing....60

Preparation of Plasmid DNA templates....61

Denaturing of Plasmid DNA....62

Priming of Denatured DNA....62

Labeling of Plasmid DNA....63

Termination 63

Chapter Page

Denaturing Polyacrylamide Sequencing Gel Electrophoresis.....64

Nucleotide Sequence Analysis and Molecular Modeling.....64

Southern Hybridization.....65

Blotting.....65

Pre-hybridization.....65

Nick Translation.....66

Hybridization.....66

In vitro Transcription and Translation.....67

E. coli S30 Extract System for Circular DNA.....67

Gel Analysis of Translation Products.....68

Discontinuous SDS-Polyacrylamide Gel Electrophoresis.....68

Fluorography.....70

Western Blot.....70

Preparation of E/ coli Lysate (for absorption of anti-*E. coli* Antibodies).....72

Absorption of Anti-*E. coli* Antibodies 73

DNA Amplification using the Polymerase Chain Reaction.....73

Protein Purification.....74

Protein Purification Using the pMal P_2 System (pilot).....74

Purification of the Target Protein Using the pET system.....76

Affinity Chromatography.....77

Processing of the Purified Sample.....78

Chapter Page

HPLC analysis....79

Extract preparation for methyltransferase assay....80

Total protein extract preparation from *S. pneumoniae*....80

Total protein extract preparation from *E. coli* XL1-Blue MRA cells....81

Preparation of Soluble protein fraction from *E. coli* expression host....81

Protein concentration determination (Bradford Assay) 82

Methylation activity assay....83

IV. RESULTS 84

Cloning of the 3. 27-kb *Eco*RI fragment in *E. coli*....85

Cloning in *S. pneumoniae*....87

Spontaneous Deletion of the cloned 3.27-kb *Eco*RI fragment in *E. coli*....88

Nucelotide Sequence Determination....90

Structural analyses of the nucleotide sequence obtained....93

Sequence similarity search....99

ORF6β Homology....109

Protein Modeling....109

In vitro Transcription and Translation....111

Construction of pSJ117....114

In vitro Assay for methyltransferase activity....116

Over expression and purification of the methylase....119

Chapter Page

The pMal-2 system....119

Cloning of ORF6α in $pMalC_2$....120

Cloning of the methylase in $pMalP_2$....121

Protein purification using pSJ136 (pilot) 122

The pET system....122

Cloning of the methylase into $pET30b^+$....124

Expression of the (ORF6α) methylase....125

Protein purification from the inclusion bodies....125

Cloning of ORF6β in an expression vector....128

Production of anti-ORF6α antibody....130

Optimization of the production of the methylase....130

Purification of the methylase from the Soluble fraction....134

Site of Methylation....148

In vivo Mutagenesis of ORF6....148

Pneumococal Transformation to check for the Em insertion....151

Southern Hybridization....153

Transformation of DP1002....156

Conjugal transfer properties of the Insertion mutant....156

Construction of pSJ119....157

Construction of SP1273....159

Transformation of SP1273....159

Conjugation of the methylase insertion mutant with *S. pneumoniae*.....160

Conjugal transfer of the methylase insertion mutant transposon to *S. pyogenes* 160

Conjugal transfer of the methylase insertion mutant Tn*5252* to S. *pneumoniae* (*Dpn*II$^+$).....162

Construction of pSJ144.....165

In vivo deletion mutation of the methylase to generate SP1213.....165

Construction of pSJ126.....168

Construction of pSJ150.....170

Construction of pSJ170.....172

Construction of pSJ172.....172

Construction of *in vivo* deletion mutant (SP1214) 174

Transformation to check if the insertion is in the transposon 174

Southern Hybridization to check the insertion.....177

Creation of SP1206 and SP1207.....177

Conjugation of the methylase deletion mutants with *S. pneumoniae* DP1004 (*Dpn*II$^-$) and DP3111 (*Dpn*II$^+$).....179

Construction of SP1208.....181

Construction of SP1209, SP1210, SP1211 and SP1212.....181

Complementation analysis.....182

Chapter Page

V. DISCUSSION.....185

REFERENCES.....193

LIST OF TABLES

Table Page

I. Streptococcal Strains....39

II. *E. coli* strains....42

III. Selective Antibiotic Concentrations....44

IV. Plasmids and Cloning vectors....47

V. Incorporation of [^{3}H] AdoMet to nonmethylated *E. coli* GM2163 DNA....117

VI. Methylase assay using the fractions collected during the purification of the methylase from the inclusion bodies....129

VII. Methylase assay using the soluble fractions obtained from the expression hosts carrying pSJ137....132

VIII. Methylase assay using the fractions collected during the purification of the methylase from the soluble fraction....136

IX. Methylase assay using the fractions collected for the second time....139

X. Methylase assay using the fractions obtained during the step gradient purification....141

XI. Methylase assay using the HPLC fractions....144

XII. Methylase assay using the HPLC fractions collected using the buffer with 0.1% Triton and 0.05% Sarkosyl....147

Table Page

XIII. Transformation of Rx1 and DP1322 to check if the Em insertion is in the transposon.....152

XIV. Transfer property of SP1262 in the presence of multiple copies of the methylase in the recipient.....158

XV. Effect of the methylase mutation on the conjugal transfer of the transposon to *S. pneumoniae*.....161

XVI. Effect of the methylase insertion mutation on the conjugal transfer of the transposon to *S. pyogenes*.....163

XVII. Conjugal transfer properties of the methylase insertion mutant to *S. pneumoniae DpnII*$^{+}$ host.....164

XVIII. Transformation experiment to check if the Em insertion is in the transposon.....167

XIX. Transformation to show that the insertion in SP1214 is in the transposon.....176

XX. Conjugal transfer frequencies of the deletion mutant to *DpnII*$^{+}$ and *DpnII*$^{-}$ hosts.....180

XXI. Complementation analysis of the deletion mutants.....183

XXII. Some properties of the three Types of Restriction-Modification systems.....187

LIST OF FIGURES

Figures Page

1. Structure of the Composite transposon tn5253 and its derivatives.....26

2. The restriction map of Tn5252, with the region of interest highlighted.....86

3. Restriction digests of the plasmids isolated from *E. coli* XL1-Blue MRA and *E. coli* C600 cells.....89

4. Restriction endonuclease map of the 3.5 kb fragment spanning the right end of Tn*5252*.....91

5. The restriction map of the 3.5-kb *Eco*RI fragment 92

6. The nucleotide sequence of the 3.5-kb right junction fragment.....94

7. Multiple alignment of ORFα with other 5^{m}C-cytosine methylases.....100

8. Alignment of variable region of ORF6α with that of the methyltransferase encoded by *Bacillus subtilis* phage SPR.....103

9. Alignment of TRDI of ORF6α with the corresponding domains of other multi-specific and mono-specific methylases.....105

10. Alignment of the TRDIII of ORF6α with the correpsonding domains of multi-specific and mono-specific methylases.....107

Figures Page

11. Alignment of a region next to TRDIII of ORF6α with the corresponding domains of multi-specific methylases.....108

12. Alignment of the predicted protein product of ORF6β with the β-subunit of M.*Eco*HK311.....110

13. Homology based modeling of the catalytic domain of ORF6α.....112

14. *In vitro* Transcription and Translation using the *E. coli* S30 extract with pLS1 and pSJ107 as templates.....113

15. Restriction digest showing the loss of the *Kpn*I site in pSJ117.....115

16. Graph showing the concentration dependence of the methylase activity.....118

17. Induction of MBP-methylase fusion protein expression from pSJ136.....123

18. Determination of the optimal expression host for expressing the methylase fusion protein by SDS-PAGE analysis.....126

19. Large scale purification of the methylase fusion protein from the insoluble fraction.....127

20. SDS-PAGE analysis of the soluble fraction obtained from pSJ137 carrying cells induced with varying amounts of IPTG.....131

21. Western Blot analysis of the soluble fraction obtained from cells carrying pSJ137, induced with 0.05 mM and 1.0 mM IPTG.....133

22. SDS-PAGE analysis of the fractions collected during the purification of the methylase from the soluble fraction.....135

Figures Page

23. SDS-PAGE analysis of the different fractions collected during the purification of the methylase from the soluble fraction for the second time.....138

24. SDS-PAGE ananlysis of the fractions collected during the methylase purification with step gradient.....140

25. UV absorbance during the purification process by HPLC.....143

26. SDS-PAGE analysis of some of the HPLC fractions.....145

27. UV absorbance of the HPLC fractions collected using an elution buffer containing 0.1% Triton and 0.05% Sarkosyl.....146

28. Agarose gel picture of the restriction digest of methylated pBR322.....149

29. Strategy for insertion mutagenesis of ORF6.....150

30. Restriction map of SP1262, the Em^r insertion mutant.....154

31. Autoradiogram showing the insertion of heterologous DNA within the ORF6.....155

32. Deletion mutagenesis strategy used to mutate ORF6.....166

33. Construction of pSJ126 using pRL425 vector and Kan^r from pDL276.....169

34.Construction of pSJ150, using pDG792 vector and Kan^r from pAT21-1.....171

35. Construction of pSJ170, using pSJ150 and pSJ126.....173

36. Restriction map of SP1213 and SP1214.....175

37. Autoradiogram showing the deletion mutation in SP1213 and SP1214.....178

CHAPTER I

INTRODUCTION

Tn*5253* (formerly called as the Ω*cat-tet* element or the BM6001 element) was identified as a part of the chromosome of the clinical isolate, *Streptococcus pneumoniae* BM6001 (176). It is a 65.5kb long site-specific conjugative transposon carrying the antibiotic resistance determinants for chloramphenicol and tetracycline (208). By inserting heterologous DNA at various sites within the transposon, Vijayakumar, *et al.* (208) have generated a detailed restriction map of the entire transposon. Earlier work indicated that Tn*5253* is a composite transposon consisting of two independent conjugative transposons, a smaller one named Tn*5251* and a larger one named Tn*5252* (7). The smaller element, Tn*5251*, carries the antibiotic resistance determinant for tetracycline and is 18-kb in size. It is a member of the Tn*916* family of transposons. Like Tn*916*, Tn*5251* can conjugally transfer from one cell to another and can integrate at multiple target sites. The point of insertion of Tn*5251* within Tn*5252* is in the middle region. Tn*5252* is a 47.5kb long conjugative transposon that carries the antibiotic resistance determinant for chloramphenicol. The genetic integration of Tn*5252* is site-specific. The unique target site of Tn*5252* in pneumococcus is identical to that of its parent element (Tn*5253*) (7). Earlier work has shown that the *cat* (chloramphenicol

acetyl transferase) gene, located in the middle of Tn*5252*, is flanked by IS-like elements. This accounts for the spontaneous curing of the *cat* gene. The observation that Tn*5251* and other IS (Insertion Sequence)-like elements integrate within the middle region of Tn*5252* seems to suggest that Tn*5252*-like elements could auto-accumulate resistance genes or other transposons. Tn*5252* is the prototype of the larger composite conjugative transposons. A functional map of the transfer-related regions of the transposon has been generated (97). The nucleotide sequences of the target, the right, and the left termini have been obtained (206). Two proteins that are involved in the transfer process have been purified and their activity identified (188, 189). An operon consisting of genes encoding transport proteins has been identified (1, 2).

Previous studies have indicated that the left and the right ends of the transposon are important for the conjugative transfer process. Since no structural or functional information of the right end of the element was available, this work was initiated to investigate the right junction region of the conjugative transposon Tn*5252*. Analysis of the 3.27-kb *Eco*RI right junction fragment of the transposon demonstrated the presence of a DNA cytosine-methyltransferase gene that has been shown to be functional. Presence of DNA methyltransferases in a variety of gene transfer systems has been reported. The role of DNA methylation in these systems will be described briefly.

(a) Regulation of Gene Expression:

Methylation has been shown to play a very important role in controlling the transposition frequency of some transposons.

Tn*10*:

Tn*10* is a compound transposon found in *Escherichia coli*. It carries the antibiotic resistance determinant for tetracycline. Tn*10* is flanked by IS*10* on both ends and the two IS elements are in an inverted orientation relative to one another. IS*10*-Right is a fully functional insertion sequence that is capable of independent transposition (157). IS*10*-R carries a transposase gene that reads from outside-to-inside. To promote transposition of Tn*10*, the transposase protein can act on the outside ends of both IS elements (IS*10*-R and L). The protein can also act on one outside end and one inside end and promote the transposition of IS*10*. To promote chromosomal rearrangements, the protein can act on two inside ends (58). Transcription of the transposase gene has been shown to be infrequent (151, 179). It has also been shown that the translation of the transposase transcript is very inefficient (178). In some *E. coli* mutants Roberts, *et al.* (157) observed an increase of 10-20 fold in Tn*10* transposition. Subsequent analysis revealed that these effects were due to mutations in the *dam* gene. There are two *dam* sites (GATC) in IS*10*, one site overlaps the -10 region of the transposase promoter, and the other is at the other end of IS*10* at the inner terminus. Methylation at the promoter region decreases the transcription of the gene. Methylation at the inner terminus decreases the ability of transposase to utilize these ends for transposition or

rearrangement. It has been shown that the promoter and the inner terminus of the IS element function coordinately (99). Absence of methylation results in increased activity. Hemimethylated species methylated at both sites on the template strand are much more active than that methylated on the non-template strand (99). These and other relevant observations suggest that IS*10* preferentially transposes when the element is in an unmethylated or a transiently hemimethylated state. Thus, the transposition of Tn*10* occurs when the replication fork just passes through or after conjugal transfer, during the complementary strand synthesis.

Tn*5*/Tn*903*:

Tn*5* is a compound transposon that carries the gene for kanamycin resistance. It is very similar to Tn*10*. Tn5 is flanked by IS*50*-R and IS*50*-L. As in Tn*10* the two IS elements are in an inverted orientation relative to one another. Tn*5* also has *dam* sites at similar positions that are regulated by *dam* methylation (222). Hemimethylation activates the transposition activity (51,134). Another element, Tn*903*, which has only one *dam* site in the promoter region of the transposase (157), has also been shown to be regulated by *dam* methylation (99).

Bacteriophage Mu *mom* gene:

The *mom* (modification of Mu) gene product modifies the adenine in the sequence 5'-(G/C)A(C/G)NPy-3' and the modification leads to the formation of N6-carboxy-methyladenine (93). The modified DNA is protected against Type I and II restriction endonucleases. This modification is toxic to the cell. Thus, the expression of this gene is very efficiently controlled and limited to a short interval late in the development of phage Mu. Expression of the *mom* gene requires the host Dam function. There are three *dam* recognition sites upstream of the *mom* promoter region (192) and methylation of these three sites stimulates the expression of *mom* gene about 20-fold (74). If one or more of the GATC sites are removed, the transcription of the *mom* gene becomes *dam* independent (146). However, Dam methylation has been shown not to be the only regulator of expression of the *mom* gene (93).

Bacteriophage P1 *cre* gene:

The *cre* gene of P1 produces a site-specific recombinase (Cre) that promotes recombination at the *loxP* sites. Cre is required in at least two stages of the P1 life cycle: (a) immediately after infection, to avoid degradation of the terminally redundant linear phage DNA. (b) After replication, for efficient partitioning of P1 monomers into the daughter cells. The *cre* gene has three relatively weak promoters, P1cre, P2cre, and P3cre. P1cre is the most upstream of the three and has two *dam* sites in its -35 region. Expression of *cre* from this promoter is increased 3-4 fold in a *dam* mutant strain (192).

P2cre has one *dam* site in its -35 region and there is no difference in expression between a wild-type *E. coli* and *dam* mutant. P3cre has no *dam* sites in its promoter region. Immediately after injection of the P1 DNA, the promoters are in an unmethylated or a hemimethylated state. During this time the *cre* gene is rapidly transcribed and Cre is made in abundance. Cre protein binds to the *lox*P sites and circularizes the phage DNA, thus avoiding degradation of the ends. Subsequent methylation decreases the extent of transcription but a low level of transcription has been shown to be enough for the partition process (192).

Bacteriophage P1 *pac* cleavage:

The packaging of bacteriophage P1 DNA into the capsid is initiated by cleavage at a unique site called the *pac* site. Packaging proceeds in one direction from the cut end until a headful of DNA has been packaged, usually more than one monomer. This headful packaging mechanism ensures a terminal redundancy. An important step in the packaging process is cleavage at one, at most, a few *pac* sites per concatemer. If each *pac* site were cut at the same time then monomers would result and there wouldn't be a terminal redundancy. The *pac* site has been localized to a 150-160 bp segment of DNA that contains seven *dam* methylation sites, three at one end and four at the other (193). Deletion of either one or both of the clusters inactivates the *pac* site. Moreover, unmethylated *pac* sites are insensitive to cleavage. Sternberg and Coulby (193) have

shown that P1 codes for its own Dam-like methylase. During the early stages of infection, the replicating DNA is not completely methylated and so the *pac* sites are unavailable for cleavage. Later in the infection some of the *pac* sites become methylated, due to the expression and action of the phage encoded methylase. As soon as some of the *pac* sites are methylated, the *pac* site is available for cleavage and the process of packaging itself blocks further cleavage of the *pac* sites in the same concatemer (193).

(b) Anti-Restriction:

A number of bacteriophages that have the ability to infect *E. coli*, *Salmonella*, *Shigella*, and *Bacillus* species, have a variety of mechanisms by which they overcome the restriction activity of the hosts. The presence of unusual bases in the genome of the infecting bacteriophages has been shown to provide resistance to restriction (15, 16, 75). Proteins that can inhibit the restriction activity of the host have also been isolated (127, 128). Bacteriophages T3 and T7 have developed an efficient mechanism for anti-restriction, they contain the gene *0.3* that codes for the Ocr protein. Ocr has been shown to bind the type-I restriction enzymes (105) and inhibit the restriction and the modification activities (10,106, 187). In addition to Ocr, T3 (but not T7) also codes for an S-adenosylmethionine (AdoMet) hydrolase (196). Thus, T3 proteins not only block type-I restriction but also inhibit type-III restriction, as AdoMet has been known to stimulate type-III restriction activity.

Phage λ also has an anti-restriction mechanism specific for type-I R/M system. The anti-restriction protein "Ral" has been shown to reduce restriction. Its primary effect has been shown to be stimulating the methylation of unmodified DNA (105), a reaction which has been known to be very inefficient in type-I R/M systems. Recent studies have confirmed that Ral stimulates the modification activity of both the 3-subunit restriction enzyme and the 2-subunit modification enzyme (124).

Sequence biases in the phage DNA have also been shown to be a mechanism by which the host-encoded restriction is avoided (21, 86, 118). This is clearly demonstrated by phage T7. The estimated number of sites for the two host-encoded Dam (GATC) and Dcm (CCA/TGG) restriction-modification (R/M) systems are 114 and 56, respectively (170). However, only six Dam sites and five Dcm sites (52) have been shown to be present in this phage. This indicates that selection was directed against these sequences as the presence of these sequences was disadvantageous for the production of phage particles. Selection against Dam sites and other restriction enzyme recognition sites in the genomes of other phages that infect members of the family *Enterobacteriaceae* have also been demonstrated (133,174). In some cases, the presence of few recognition sites of certain restriction enzymes can be tolerated as they do not lead to restriction. For example, in the phages T3 and T7 there are very few *Eco*RII sites and these are not restricted. The restriction enzyme requires at least two sites in close proximity for restriction and in the case of T3 and T7 they are far away for any restriction (107).

Another mechanism by which phages avoid host-encoded restriction is by specifically modifying certain bases with the help of modification methylases that they

encode. This increases the ability of the phage to infect the host effectively as the modified base makes the DNA resistant to cleavage by the host-encoded restriction enzyme. A variety of temperate phages of *Bacillus* code for such modifying enzymes. These enzymes have been well characterized. If the phages have to infect a host effectively then they have to modify the bases that the new host's restriction endonuclease would recognize. To accomplish this, the phages of *Bacillus* code for multi-specific methylases that recognize more than one target site and modify the cytosines in them. For example, the *Bacillus amyloliquifaciens* phage H2, which codes for a methylase (M.H2II), methylates the target sequence GGATCC at the 5'-cytosine residue, which is also the recognition sequence of the host-encoded restriction system (*Bam*HI) (43). The methylation by the *B. subtilis* phage ϕ3TII- and ρ11s encoded methylases that modify TCGA overcomes the restriction potential of the *Bsu*MI system (PyTCGAPu) (22, 88). The methylation of GGCC by different phages (SPR, Spβ, ϕ3TI, ρ11s, ρ11b, and H2) prevents the restriction by *Bsu*RI (GGCC). The SPR phage-encoded methylase (M.SPR) modifies the 3' cytosine in CCGG, and this modification protects the sequence from *Bsu*F1 restriction (CCGG). There are other methylation specificities identified in the suite of phage methylases that correspond to no known host restriction-modification specificities (143). It is very likely that these specificities are directed towards other phage sensitive bacteria that might have as yet unidentified restriction-modification systems that match these specificities.

Hill, *et al.* (77) have demonstrated the presence of a DNA methyltransferase gene in the genome of *L. lactis* phage ϕ50. They have also shown that this was probably

picked up recently from a plasmid as a 1.3-kb of DNA sequence of the phage coding for the methylase was identical to that of the plasmid. Conjugative plasmids in the incompatibility groups I and N have also been shown to code for anti-restriction functions. This anti-restriction activity coded by the *ard* genes has been shown to be directed towards type-I R/M systems. The *ard* genes were located near the origin of transfer and they are expressed very early in the conjugation process in the recipient (12, 13, 49, 55, 154).

The presence of a DNA C5-cytosine methyltransferase in Tn*5252* could account for an observation made by Guild, *et al.* (64), several years ago. They observed that the transfer of a plasmid, pIP501, from Rx1, a null strain lacking both the restriction and modification genes of *Dpn*II system (for *Diplococcus pneumoniae*) was restricted in a *Dpn*II$^+$ recipient (as compared to a *Dpn*II$^-$ recipient) whereas the transfer of the Tn*5253*, from the same donor cell to the *Dpn*II$^+$ host was not. However, the chromosomal DNA isolated from the host harboring Tn5253 was sensitive to *Dpn*II. This suggested that the transposon was not subjected to restriction during conjugation. This lack of restriction of the transposon was interpreted to mean either one of the following two things: (a) during the mating, only one strand is transferred or (b) the DNA is in a protected state during the transfer process.

As mentioned previously, bacteriophages and some conjugative plasmids have been found to carry methyltransferase genes to protect their DNA from the host-encoded restriction systems during infection or transfer. Although transposition of some transposons have been shown to be controlled by host-encoded methylases. Tn5252 is

the only transposon shown to encode a methylase. A thorough investigation of this methyltransferase would help us understand if this gene is involved in regulating the conjugal transfer process and determine if it is responsible for the anti-restriction reported by Guild, *et al.* (64).

Specific aims

1. To obtain the sequence of the right junction 3.27 kb *Eco*RI fragment of DNA between coordinates 46.4 and 49.7 of Tn*5252* and study the structural organization of this region.

2. To analyze the DNA sequence for open reading frames (ORFs), sequence similarities, and identify the role(s) of one or more of the genes by site-specific mutagenesis.

3. To purify the gene product(s) of one or more of the genes for structural and functional studies.

CHAPTER II

LITERATURE REVIEW

The Organism: *Streptococcus pneumoniae*

S. pneumoniae is a Gram positive, lancet-shaped diplococcus that belongs to the family *Streptococcaceae*. It is non-sporeforming and non-motile. *S. pneumoniae* is a facultative anaerobe and the main product of its metabolism is lactic acid. It is catalase and peroxidase negative and has very complex nutritional requirements. This organism is carried in the upper respiratory tract of many healthy individuals. The immune status of the individual at the time of colonization and the virulence of the infecting strain determine whether the strain remains confined to the nasopharynx or becomes invasive.

S. pneumoniae was first isolated in 1881 from human saliva by Pasteur and Sternberg in independent studies (91). Since then, it has been extensively studied. These studies led to the important scientific discoveries of Avery, *et al.* in 1944 (5). In spite of extensive research on this organism, a comprehensive list of all its virulence factors is far from complete (3, 90).

S. pneumoniae still remains one of the leading causative agents of bacterial pneumonia and bacteremia. It is the second most frequent causative agent of bacterial meningitis (46) and otitis-media (141), after *Haemophilus influenzae* type b. The incidence of pneumococcal infections is highest in infants under two years of age and in people over 60 years of age. Each year, in developing countries a large number of

children under the age of five years die due to pneumococcal pneumonia. In the last four decades mortality due to pneumococcal bacteremia has remained stable between 25 and 29% of all the deaths due to dieseases (3) in spite of the availability of antibiotics.

Drug Resistance in *S. pneumoniae*

Pneumococcus remained susceptible to most of the antibiotics during the early post-antibiotic era, in spite of isolation of a sulfonamide-resistant strain as early as 1939 (70). Sulfonamide resistance was found to be linked to mutations at a chromosomal locus (81, 82). Tetracycline resistance appeared in clinical isolates in 1963 (57), followed by penicillin resistance (72), and resistance to MLS (Macrolide, Lincosamide and StreptoGramin) antibiotics (50) in 1967. Further investigations revealed that penicillin resistance was not due to the production of β-lactamase (159) but due to accumulation of chromosomal mutations that confer stepwise resistance (223). Resistance to MLS antibiotics was found to be due to the *erm* gene whose product methylates the 23S subunit of the ribosomal RNA (213). Chloramphenicol resistance, due to the production of chloramphenicol acetyltransferase (which detoxifies chloramphenicol), was reported in 1970 (47, 138). In 1977, kanamycin resistance, due to the production of aminoglycoside 3'-phosphotransferase was reported (26). In the 1970s, there was a dramatic appearance of multiply-resistant pneumococci in South Africa (87), Paris (47) and Japan (138). All these clinical isolates were resistant to chloramphenicol and tetracycline at very high levels. One such strain isolated from a patient suffering from sinusitis in Paris was the strain BM6001 (47).

Source of Antibiotic Resistance

Antibiotic resistances have often been shown to be carried on R-plasmids. Several observations led to the possibility that Cm^r, Tc^r antibiotic resistances of strain BM6001 could also be due to the presence of such plasmids. The observations were:

(i) Chloramphenicol, erythromycin and kanamycin resistance determinants of clinical isolates of pneumococcus were similar to the plasmid-encoded resistance determinants of Staphylococci and other Streptococci (181).

(ii) BM6001 and the strain from Japan, N77, both were found to contain inducible chloramphenicol acetyltransferase activity like the ones often associated with plasmids.

(iii) Dang-Van, *et al.* (47) found that following growth in ethidium bromide medium, BM6001 segregated chloramphenicol-sensitive variants at a frequency approaching 10^{-3} CFU/ml , but these remained Tc^r.

(iv) The DNA from BM6001 transformed laboratory strains to chloramphenicol resistance only at a very low frequency and did not transform to tetracycline resistance at all.

Dang-Van, *et al.* (47) and Robins-Brown, *et al.* (159) attempted to isolate covalently-closed-circular molecules from these clinical isolates but could not demonstrate the presence of plasmid DNA in these isolates. Smith and Guild (182) also were not able to detect any plasmid DNA from the strain BM6001 while they were able to isolate a cryptic, 2.0 megadalton plasmid, pDP1, from another virulent strain of pneumococcus.

As bacteriophages have been demonstrated in *S. pneumoniae* (14) and in the BM6001, Dang-Van, *et al.* (47) investigated whether the chloramphenicol and

tetracycline resistances in BM6001 were associated with one or more prophages. Analysis of 100 independent R6 clones lysogenized by the phages from BM6001 for resistances to chloramphenicol or tetracycline or both showed that all clones were susceptible to both the antibiotics like their parent strain, R6. This eliminated the possibility of phage borne antibiotic resistance (47). Thus, the absence of plasmid DNA and phage borne antibiotic resistance led to the possibility that these antibiotic resistances might be chromosome- borne.

Cm^r and Tc^r are Chromosomal Insertions

The first convincing evidence that chloramphenicol and tetracycline resistances were due to heterologous insertions in the chromosome came from genetic and physical data. Shoemaker, *et al.* (176) observed that (i) transformation of Rx1 (wild-type *S. pneumoniae*) with the DNA from BM6001 gave rise to Cm^r and Tc^r colonies at frequencies that were less than 1% of those expected for transformation by homologous point markers. They also observed that the transfer frequency of the *tet* gene was eight-fold lower than the transfer frequency of the *cat* gene. (ii) Transfer of *tet* was extremely sensitive to shearing of the DNA, whereas the transfer of *cat* was less so. (iii) the transfer frequency of *cat* to a strain that was Tc^r, DP1320, was observed to be 30-fold higher than that of Rx1. Similarly, the presence of *cat* in the recipient, DP1321, increased the receptivity for *tet* by two-four fold. (iv) *cat* activity was observed to co-sediment with the chromosomal marker (*thy*) and velocity distribution of *cat* was found to shift by the shear treatment of DNA. (v) In dye-buoyant density gradients, *cat* activity

was found to co-band with the chromosomal marker (*nov-1*) at the position expected for linear DNA and not with that of plasmid DNA.

(vi) *cat* was found to be linked to *nov-1*, a chromosomal marker. (vii) *tet* was also linked to *cat*. Thus, it was concluded that *cat* and *tet* loci are a part of a large heterologous insertion in the chromosome of *S. pneumoniae* BM6001. Shoemaker, *et al.* (176), estimated the lengths of *cat* region to be approximately 4-8 kilobases (kb) long and that of *tet* to be above 30 kb. This heterologous chromosomal insertion carrying the *cat* and *tet* region was called Ω*cat-tet.*

Ω*cat-tet* is Conjugative Transposon

Shoemaker, *et al.* (177) reported that Ω*cat-tet* element could be transferred from one cell to another, through a process that requires cell-to-cell contact. The frequencies of this transfer ranged from 10^{-6} to 10^{-5}/donor, unlike the transfer of plasmids which were in the range of 10^{-3} to 10^{-1}/donor (185). The transfer process was *rec*A independent and DNase resistant, no mobilization of chromosomal markers was observed. The transconjugant that now contains the Ω*cat-tet also* behaved like the parental strain. They also observed that the transfer took place only on solid surfaces, like nitrocellulose filters, and not in broth (177, 183). Both intra- and inter-specific conjugative transfers of Ω*cat tet* were observed (26, 80). Thus, these Ω*cat-tet* insertions and other similar elements detected in other species (26, 47,59, 65, 83) were termed "Conjugative Transposons".

Uniqueness of Conjugative Transposons

Conjugative transposons are unique. They differ from other gene transfer systems such as phages, plasmids, and transposons, but, at the same time, share some features with them. Conjugative transposons are phage-like. Their integration and excision from the host DNA (plasmid or genomic) resemble that of the lambdoid phages (161). In fact, the nucleotide sequences of the genes involved in the integration and excision processes reveal that they code for enzymes of the lambda integrase family (97, 149, 150, 153, 197). However, these transposons do not form infectious particles.

The conjugative transposons are plasmid-like. They are transferred by conjugation and have a covalently-closed-circular intermediate during their transfer. However, these circular intermediates do not replicate like plasmids, at least in the hosts tested so far (156, 173).

The conjugative transposons are transposon-like, as they excise from, and integrate into DNA. However, their mechanisms of integration and excision are different from those of the well-studied transposons like Tn*5* and Tn*10* (38, 156, 161, 171). Unlike transposons, these elements can mobilize themselves from one cell to another by the process of conjugation.

Classification of Conjugative Transposons

Conjugative Transposons were first identified in Gram-positive bacteria but now a variety of them have been identified in Gram-negative bacteria. They differ from one

another in their size, the traget specificity, the antibiotic resistance they carry, the host range etc.

Conjugative Transposons of Gram Negative Bacteria:

Conjugative Transposons have been identified in several species belonging to the genera *Neisseria*, *Kingella*, *Vibrio*, *Salmonella* and *Bacteroides*. Tn*916*-type of conjugative transposons have been observed in *Neisseria gonorrhoeae* and in *Kingella* spp (103, 158, 200). As Tn*916* can mobilize itself between Gram-positive, as well as, Gram-negative bacteria, this observation is not all-together surprising. Waldor, *et al.* (210) have identified a novel site-specific conjugative transposon (named SXT) containing the determinants for sulfamethoxazole, trimethoprim and streptomycin resistance in *Vibrio cholerae* O139. SXT is a 62-kb transposon, that has been shown to mobilize itself from *V. cholerae* O139 to *E. coli*. It has also been shown that SXT can also transfer itself from one species to another (210). Ctnscr*94*, an orientation-specific conjugative transposon, was identified in *Salmonella senftenberg* 5494-57 (218). This transposon has been shown to insert at two specific sites in the chromosome. The transposon is 100 kb in size (78) and codes for the enzymes needed for sucrose fermentation. It is the first conjugative transposon identified in enteric bacteria.

The best studied among the Gram-negative conjugative transposons are the non-replicating Bacteroidal Units (NBUs). This group of transposons has been identified in Gram-negative anaerobic *Bacteroides* species (161). These transpsons range in size from 65 to over 150 kb (162). Most of the NBUs have been shown to carry a resistance

gene for tetracycline of the *tet*Q typeand not the *tet*M type carried by Tn*916*. Unlike other conjugative transposons, these elements have been shown to integrate into co-resident plasmids and mobilize them in *cis*. They have also been shown to mobilize other unlinked integrated transposons (161, 162).

Conjugative Transposons of Gram Positive Bacteria:

Two transposons that belong to this category were the first conjugative transpsons to be identified. They are Tn*916* and Tn*5253*. The conjugative transposons of Gram-positive bacteria can be classified into two classes based on their sizes and sequence similarity:

(a) Tn*916*/Tn*1545* family

(b) Tn*5253* family.

(a) Tn*916*/Tn*1545* family:

These are smaller conjugative transposons whose size ranges from 18 kb to 25 kb. The transposons that belong to this family usually carry the *tet*M-type determinant conferring tetracycline resistance. The *tet*M marker provides the host with the ability to resist tetracycline as well as minocycline, a semi-synthetic derivative of tetracycline (172).

Tn*916*:

Tn*916* was one of the first conjugative transposons identified. It was found in the chromosome of *Enterococcus faecalis* DS16 (59). It carries the *tet*M determinant for tetracycline resistance and is a "simple transposon", in the sense that it transposes as a single unit (172). Tn*916* has been shown to insert into various sites on a highly conjugative plasmid, pAD1, and also into various sites on the chromosome of recipient bacteria. The transfer process has been shown to be *rec*A independent. The transconjugants carrying independently-acquired Tn*916* have widely different donor potentials for subsequent transfer. Transfer frequencies range from $<10^{-9}$ to $>10^{-4}$ per donor. These transconjugants have also been shown to differ in the number of copies of Tn*916* they acquired, half the members of a population of transconjugants have more than one copy of Tn*916*. The integration event of Tn*916* has been shown not to generate direct repeats at its junctions (27, 42). The presence of either four or five T-residues at the end of the transposon has been observed after the transposition of Tn*916*. This has been suggested to be related to the slippage at the right end of the element by the site-recognition apparatus (38). An excised circular form of Tn*916* has been isolated from *E.coli*. This circular form has been shown to integrate into different chromosomal locations upon transformation into *B. subtilis* protoplasts (173), thus providing physical evidence for the presence and involvement of a circular intermediate in transposition. This transposon has been studied extensively and its complete nucleotide sequence has been determined recently (41).

Tn*1545*:

S. pneumoniae BM4200 is a multiple antibiotic-resistant clinical isolate. It has been shown to be resistant to sulfonamides, trimethoprim, penicillins, chloramphenicol (*cat),* macrolide-lincosamide-streptoGramin B-type (MLS) antibiotics (*erm*AM), tetracycline (*tet*M) and kanamycin (*aph*A) and other structurally related aminoglycosides (26). Buu-Hoi, *et al.* and Smith, *et al.* (26, 65) have shown that all the resistances were chromosome borne. They have also observed the conjugative transfer of *cat-tet-erm-aph*A but they did not observe the transfer of any of the other resistances. Carlier and Courvalin (29) cloned the *tet-erm-aph*A determinants in *E.coli* and observed that the segment exhibited *rec*A-independent transposition and named this element Tn*1545*. This transposon is 25.3 kb in size and the transposition and the structural features of this element have been shown to be similar to that of Tn*916*.

The Tn*916*/Tn*1545* family of transposons includes members that closely resemble Tn*916* and so far several members of the family have been identified. Some of the family members include, Tn*919* from *S. sanguis* (38); Tn*918*, Tn*920*, Tn*925* and Tn*3702* from *E. faecalis* (38, 39); Tn*3703* from *S. pyogenes* (120); Tn*5031*, Tn*5032* and Tn*5033* from *E. faecium* (38); and Tn*5251* from *S. pneumoniae* (7).

The host range of this family of transposons has been observed to be remarkably broad. Members of this family have either been isolated from or have been introduced into at least 52 different species (both Gram-positive and Gram-negative) from 24 genera (41).

(b) Tn*5253* family:

The members of this family of conjugative transposons are larger in size than the Tn*916*/Tn*1545* family of transposons. They are above 50 kb in size and they usually code for multiple antibiotic resistances. Tn*5276*, found in the chromosome of *Lactococcus lactis* NIZO R5, Tn*3701*, found in the chromosome of *S. pyogenes* A454, and Tn*3951*, found in the chromosome of *S. agalactiae* B109 are all members of this family.

Tn*5276*:

This conjugative transposon is 70 kb in size. It has been shown to carry the gene necessary for the production of nisin, a lantibiotic that is used to prevent the growth of undesirable bacteria in the dairy industry. It also has been shown to carry the genes that are needed for sucrose metabolism, for reduced bacteriophage sensitivity and immunity to nisin (152, 153). Raush, *et al.* (152, 153) have shown that this element integrated itself into the chromosome of *Lactococcus lactis* without any target duplication. The element has been observed to have at least five different target sites. It has been shown to insert in one or more of these five sites, in an orientation-specific manner. Another nisin-sucrose element, Tn*5301* that is similar to Tn*5276* has been identified in *L. lactis* FI5876 (79).

Tn*3701*:

Tn*3701* is a composite conjugative transposon identified originally in the chromosome of *S. pyogenes* A454 (120). It is 50 kb in size and carries the determinants for chloramphenicol (*cat*), erythromycin (*erm*) and tetracycline (*tet*M) resistance. By DNA-DNA hybridization experiments, Le Bouguenec, *et al.* (120), have shown that Tn*3701* contained a 19.7 kb segment within itself that was similar to Tn*916*. This region has been shown to carry the *erm-tet* genes. They have also shown that this region could transpose itself to different sites on the chromosome from plasmids but lacked the conjugative properties. This small element within Tn*3701* has been named Tn*3703*. Le Bouguenec, *et al.* (120) have also shown that the *erm* gene and some adjacent DNA can spontaneously cure itself suggesting that the *erm* gene might be located in another transposon. The *erm* insertion into a gene that is needed for the conjugation function of Tn*3703* could have rendered that gene inactive and was predicted to be the reason for the lack of conjugative functions of Tn*3703* (119).

Tn*3951*:

This 67-kb composite conjugative transposon was found in the chromosome of *S. agalactiae* B109. It carries the determinants for chloramphenicol, tetracycline, and erythromycin resistances (84). This element has been shown to be able to conjugally transfer to other Streptococci belonging to Groups A, B, C, D, G, and H (80, 83). The genetic integration of this element has been shown to be site-specific (84). Smith and

Guild (184) demonstrated that this element can transpose into pAD1, a hemolysin plasmid of *E. faecalis*. The *erm-tet* determinants have been shown to reside in a 22 kb *Eco*RI fragment (84) and this segment has been shown to be similar to Tn*916*, like Tn*3703* (120). Inamine and Burdett (84) have also shown that this 22-kb *Eco*RI fragment, when cloned in *E.coli* DB11, expressed both *erm* and *tet* resistances.

Tn*5253*:

As mentioned earlier, this element was originally identified in the chromosome of *S. pneumoniae* BM6001 and it carries the determinants for chloramphenicol and tetracycline resistance. This element was referred to as the BM6001 element or the Ω*cat-tet* element and later named Tn*5253*. Tn*5253* is the most extensively studied member of this family. This is a composite conjugative transposon whose integration is site-specific. This transposon has been shown to transfer to other streptococci during filter-matings at a frequency of 10^{-7} to 10^{-6} per donor (181). Vijayakumar, *et al.* (208) have constructed a restriction map of this 65.5 kb transposon by inserting an *E. coli* vector, pVA891, which is incapable of autonomous replication in Streptococci, at a number of sites specifically within the transposon. This enabled them to identify the junctions and the target site of the transposon (207). DNA-DNA hybridization of Tn*5253* has shown that the DNA segment carrying the *tet* determinant was similar to Tn*916*, like Tn*3703* and Tn*3951* (120). This observation showed that there is a region in Tn*5253* that is homologous to Tn*916* but does not answer the question of whether it is capable of transferring itself independently like Tn*916* or not (as in the case of

Tn*3703*). Ayoubi, *et al.* (7) cloned a 23-kb *Xba*I fragment of Tn*5253* into the pVA891 vector to generate the plasmid, pVJ403, and introduced it into *E. coli*. In the absence of selection (tetracycline), they observed that there was a deletion of about 18-kb from this plasmid. When pVJ403 was introduced into *S. pneumoniae*, it was observed that the element carrying the *tet* gene (18-kb) transposed into the chromosome. This indicated that, like Tn*3703*, the region of Tn*5253* that was similar to Tn*916* was capable of independent transposition. It was also shown that this 18-kb segment was able to conjugally transfer itself to other streptococci and integrate at different sites in the chromosome, unlike Tn*3703*. The segment of Tn*5253* that carries the *tet* gene and can transpose independently was named Tn*5251*. Tn*5251* belongs to the Tn*916*/Tn*1545* family of conjugative elements. The remaining DNA after the excision of Tn*5251* has also been shown to transfer conjugally to other streptococci (7). This segment was named Tn*5252*. Tn*5252* is 47.5-kb long, carries the *cat* gene, and is a site-specific transposon having the same target-site specificityas its parental element, Tn*5253*. The *cat* gene has been shown to be flanked by IS-like elements (207). This accounts for its spontaneous loss at detectable frequency. These observations suggest that smaller elements like Tn*5251* were added later to larger ones like Tn*5252* to generate the composite elements such as Tn*5253* (Fig. 1). Thus, Tn*5252*-like elements should be considered the prototype of composite conjugative transposons such as Tn*5253*.

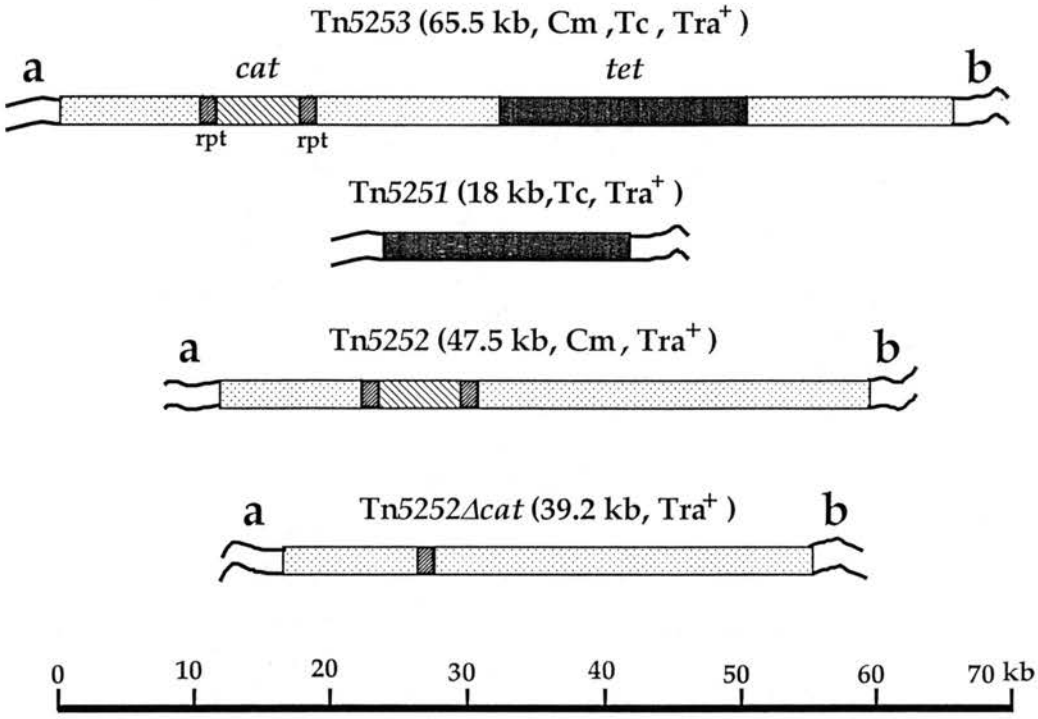


Figure 1. Structure of the Composite transposon Tn*5253* and its derivatives. The DNA remaining after the transposition of Tn*5251* is termed Tn*5252*. Spontaneous loss of *cat* occurs at a detectable frequency (because of the flanking repeats), Tn*5252* Δ*cat* is still transfer proficient. (Picture printed with permission from Kilic, *et al.* (97)).

Tn*5252*:

Unlike the Tn*916*/Tn*1545* family of transposons, these elements have never been isolated from Gram-negative bacteria and even in the Gram-positive bacteria, of these large elements are not common. As the molecular details about these elements are very limited, the focus of our lab has been to gain a broader understanding of the biology Tn*5252*-like elements. The Tn*916*/Tn*1545* family and the Tn*5252*-like transposons differ mainly in their target specificity. The integration of Tn*5252* is at a unique target site in *S. pneumoniae* (7) whereas the Tn*916*/Tn*1545* family of transposons tends to integrate at multiple sites in the chromosome (37, 45). To gain insight into the nature of integration, Vijayakumar, *et al.* (206) obtained the nucleotide sequence of the termini of the transposon (Tn*5252*) from pneumococcal transconjugants. Comparison of the sequences revealed a nearly perfect identity (with the exception of two bases) between the left and right termini (over a 72 bp sequence). The 72 bp sequence present in the left termini and the right termini have been shown to belong to the transposon and the chromosome of *S. pneumoniae* respectively. Thus, the 72 bp sequence was identified to be the signal for the site-specific integration of the element. Analysis of *S. gordonii* transconjugant revealed the presence of the 72 bp sequence in the left terminus, as expected but at the right terminus only a 35 bp sequence was observed (6). Thus, a 35 bp sequence was shown to be adequate for the site-specific integration of the transposon. To localize the transfer-related regions in the transposon, Kilic, *et al.* (97) inserted heterologous DNA (pVA891) at various sites in the transposon to create mutants. Analysis of the results of filter-mating experiments done with these mutants as donors

led to the localization of the regions of the transposon that are involved in transfer functions. Interestingly, insertions and deletions in the middle of Tn*5252* were found not to affect the transfer functions whereas the mutations at the ends were found to reduce the transfer frequency dramatically. This result was consistent with the observation that the smaller conjugative transposon, Tn*5251* and the *cat* gene flanked by IS-like elements integrated within the middle region of Tn*5252* (97). To understand the functional role of transfer-related regions of Tn*5252*, Srinivas, *et al.* (188, 189) obtained about 6 kb nucleotide sequence from the left end of the transposon. Further analysis led to the purification of a DNA relaxase and a regulatory protien. The DNA relaxase has been shown to nick in a site-specific manner within a region in the middle of the transposon (188). The regulator has been shown to bind to sequences upstream of an integrase-like gene, the DNA relaxase, another functionally unidentified gene and also to its own promoter (189). Nucleotide sequence was also obtained from another transfer-related region, located near the right end of the transposon (1, 2). Analysis of the sequence revealed the presence of an operon-like structure consisting of at least ten genes. The predicted proteins of these genes showed a tendency to form trans-membrane segments (1) and these also showed significant similarities to transport proteins from other systems. These observations suggested that these proteins might be involved in transferring Tn*5252* across the membrane (1).

As the ends of the transposon have been established to be important for the transfer, the work described in this thesis focuses on the right-most end of the transposon Tn*5252* where no structural or functional information was available.

Preliminary work to this end revealed the presence of a functional DNA methyltransferase (164).

DNA Methyltransferases

DNA methylation is a major form of modification in bacteria, fungi, phages, animal viruses, plants, and vertebrates. In prokaryotes, DNA methylation has been shown to be involved in restriction-modification (R/M), mismatch repair and gene regulation (131, 143, 203). DNA methyltransferases have been classified into three groups, Type-I, II and III based on their structure (refer Table XXII for a summary).

Type-I

Type-I is the most complex of all the three methyltransferase groups. Previous studies have indicated that these methylases are multi-functional enzymes, catalyzing both restriction and modification reactions. They have also been shown to be restriction-dependent ATPases and DNA topoisomerases (17). Magnesium, ATP, and S-adenosyl L-methionine (AdoMet) are required for restriction and modification. The restriction of DNA has been shown to be random and is far from the asymmetrical recognition sites. A separate modification methylase has been isolated recently (117) and has been shown to be an efficient mono-functional methylase, thus, disproving previous studies. Thirteen different type-I systems that had been described by the early 1990s, eight of which were found in *E. coli* (*Eco*AI, *Eco*BI, *Eco*DI, *Eco*DXXI, *Eco*EI, *Eco*KI, *Eco*R124I and *Eco*R124/3I), four in *Salmonella* species (*Sty*SBI, *Sty*SPI, *Sty*SQI and *Sty*SJI) and one in

Citrobacter (*Cfr*AI) (215). Type-I R/M systems from other bacteria like *Pasteurella hemolytica* (76), *Klebsiella pneumoniae* (123), *Mycoplasma pulmonis* (180), *L. lactis* (169) and another *Salmonella* species (*S. enterica*) (201) have been reported recently. Among the type-I R/M systems, the two methylases that have been well characterized are from *E. coli* B (117) and K12 (199). It was observed that they both contain two non-identical sub-units that are the same as two of the three sub-units of the corresponding restriction enzymes. Later analysis of the type-I R/M systems has shown the involvement of three genes in type-I systems and they are *hsd*R, *hsd*M, and *hsd*S (hsd, for host specificity for DNA). It has been shown that *hsd*M and *hsd*S are necessary for modification of the target (199) and the *hsd*S gene product has been shown to be responsible for recognizing specific sequences in DNA for both restriction and modification. In *Eco*K and *Eco*B, hemimethylated DNA has been shown by far to be the preferred substrate for methylation. Vovis, *et al.* (209) and Burckhardt, *et al.* (24) observed that methylation of a hemimethylated substrate occurs 100 times faster than methylation of a completely unmodified substrate. This observation is physiologically relevant, as hemimethylated DNA is the normal product of DNA replication or repair. It has also been shown that *Eco*B and *Eco*K are allelic and they share homology with other restriction-modification systems from *Salmonella* species (140). Recently Meister, *et al.* (136) have shown that the *hsd*S from the *Eco*DXXI system recognizes a different target site because of a Tn*5* transposition within the *hsd*S gene.

Type-III

The restriction enzymes of type-III R/M systems have been shown to be comprised of two non-identical sub-units that require ATP and Magnesium for activity. Unlike type-I enzymes, restriction is independent of the presence of AdoMet, although it stimulates the reaction. In the presence of ATP and AdoMet, restriction and modification are competing reactions (69). Restriction of DNA has been shown to occur at a fixed distance from the asymmetrical recognition sites. A separate modification enzyme lacking the endonuclease activity has been isolated from the type-III restriction-modification (R/M) system (20). Hadi, *et al.* (69) have shown that the modification methylase consists of one of the two sub-units of the restriction enzyme. This sub-unit is the one that recognizes the target sequence in both the restriction and the modification reactions. *Eco*P1, *Eco*P15, *Hin*fIII, and *Sty*LTI are the only four type-III R/M systems that have been identified so far and all four of them have been shown to methylate adenine residues in one strand only. In the case of *Eco*P1 and *Eco*P15, the recognition sequences have been shown to be AGACC (8) and CAGCAG (68) respectively. Only one strand is methylated, as adenines are present in only one strand. Both *Hin*fIII and *Hin*eI (isoschizomer of *Hin*fIII) have been shown to recognize the target sequence, CGAAT (145). Both the strands have adenine residues but only one strand is methylated. In cells harboring the type-III R/M systems, DNA replication generates one daughter DNA molecule containing the parental modification and a second daughter molecule with the corresponding site completely unmodified. This unmodified DNA should be restricted but it is not. The reason for this lack of restriction was understood

from the observation that T3 phage DNA is restricted by the *Eco*P15 R/M system whereas the DNA of its close relative T7 is not. Although T7 has 36 *Eco*P15 sites in its genome (52), all the *Eco*P15 sites (CAGCAG) in T7 are in one DNA strand and all the CTGCTG are in the other strand. In T3 (from the partial sequence that was available) the *Eco*P15 sites are in both the orientations. This led to the hypothesis that two sites have to be in an inverse orientation for restriction, and sites present in only one orientation cannot be restricted but can be modified. This hypothesis was later tested and proved to be correct (135).

Type-II

Type-II R/M systems have a simple sub-unit structure. There are separate enzymes that catalyze restriction and modification. Magnesium is required for restriction and AdoMet is required for modification. The restriction enzymes have been shown to cut at, near, the sequences that they recognize. In case of methyltransferases, only one of the strands has been shown to be methylated during each methylation event. This observation is consistent with the fact that the natural substrate for the enzyme within the cell is newly replicated hemimethylated DNA (33). Although both the endonuclease and the cognate methylase recognize the same target sequence, they have not been observed to show significant similarities in their primary sequences (30, 221). This suggests that they both have evolved independently and use different strategies to interact with identical targets (203). The lack of similarity between the endonucleases and their cognate methylases could be explained by the fact that they both catalyze different

reactions. No detectable similarities have been observed between endonucleases that have been isolated from different organisms but recognize the same target site (Isoschizomers) and cut at the same position, as in the case of *Bsu*RI and *Ngo*PII (98, 198). However, there are a few cases of isoschizomeric enzymes that have been observed to share some similarity. These pairs are, *Eco*RI and *Rsr*I (191), *Bsu*BI and *Pst*I (221), *Tth*HB81 and *Taq*I (11), *Fnu*DI and *Ngo*PII, and *Cfr*9I and *Xma*I (217).

Unlike the endonucleases, the methyltransferases share sequence similarities. Methyltransferases from a variety of organisms have been cloned and sequenced (216) and comparison of the amino acid sequences of these methylases have helped distinguish between the conserved regions and the variable regions (100,116,147). Methyltransferases have been divided into two classes based on the modification that they introduce. The first class includes enzymes that methylate the pyrimidine ring carbon yielding C5-methylcytosine. Members of the second class methylate the exocyclic nitrogens and form either N4-methylcytosine or N6-methyladenine. Two families of C5-cytosine methylases (5^{m}C-cytosine methylases) have been described. The first family includes members that are predominantly accompanied by their cognate endonucleases as a part of the R/M system. These methylases have been shown to be mono-specific in that they recognize and modify a single DNA recognition sequence (216, 101). The second family contains methylases that have been shown to be multi-specific in that these methylases recognize and methylate several different DNA sequences. These are not accompanied by any restriction endonucleases. There are very few members in this family and have so far been limited to the methylases encoded by the bacteriophages of *Bacillus*. Although the two families differ in their ability to recognize

and methylate different DNA sequences, they still have been observed to show an overall similarity except for the variable region. Analysis of several 5^{m}C-cytosine methylases have shown that they consist of ten conserved blocks (I to X) of amino acid residues, in a sequential order. Most of the conserved blocks have been observed to be separated from each other by short regions of similar length, but with no sequence conservation. Clusters of invariant positions enclose about 200 amino acids in the N-terminal half and about 50 amino acids in the C-terminal half (Fig. 7, 116, 147). The preservation of the ten conserved blocks in all the 5^{m}C methyltransferases suggest that they are responsible for functional or structural properties that are common to all the cytosine methylases. Binding to AdoMet and covalent binding to the C6-position of the cytosine to be methylated are two common functions and these two functions have been attributed to two distinct blocks. Block I (Fig. 7) consisting of F-X-G-X-G (or a similar glycine rich sequence) was presumed to be a part of the AdoMet binding site. The presence of this block was observed in a wide variety of AdoMet-dependent methyltransferases (Mtases) including the DNA, RNA, protein, and small molecule Mtases (85, 92, 100, 217). Wu and Santi (219) proposed that the 5^{m}C methylases function catalytically analogous to thymidylate synthase (166) and tRNA (uracil-5)-Mtase (165), enzymes that transfer single methyl groups from AdoMet to the C5 position of pyrimidines. A covalent intermediate has been suggested to be generated by a nucleophilic attack on the C6-position of the cytosine, by a cysteine residue, (32, 60) present in the Pro-Cys dipeptide (which is absolutely conserved in all the 5^{m}C-Mtases) of conserved block IV (Fig. 7). This has been suggested to activate the C5 position of the cytosine and this activated C5 initiates a nucleophilic-attack on the methyl group of AdoMet. Chen, *et al.* (32), and

Friedman and Ansari (60) have recently identified the sulfhydryl group in the Pro-Cys dipeptide as the active nucleophile in M.*Hae*III and M.*Eco*RII. The activity of several methylases have been shown to be abolished when this conserved cysteine is replaced (137, 214, 220). Between the conserved block VIII and IX is the variable region that had been proposed to carry the target-recognizing domains (TRDs), which are responsible for the enzymes' characteristic capacities to interact with defined DNA sequences (Fig. 7). Experimental proof for this assumption has been provided initially in the case of multi-specific methylase (9, 204, 214) and recently in the case of mono-specific methylases, M.*Hpa*II and M.*Hha*I (101). In the case of mono-specific methylases, the variable region has been observed to be 80-120 amino acids long and in the case of multi-specific methylases, it has been observed to be 200-300 amino acids long. The TRDs have been shown to represent contiguous segments consisting of about 40 amino acids each. They have been shown to be arranged consecutively without any overlap or separation by linker amino acids (203). Walter, *et al.* (212) have also shown that the TRDs can be rearranged within and exchanged among multi-specific methylases without functional loss. The methylases encoded by the bacteriophages, ρ11s and H2II have been shown to contain TRDs that resemble active TRDs of other methylases but do not contribute to the enzymes' methylation potentials (112, 113). Lange, *et al.* (112) have functionally activated these inert TRDs (also termed as Pseudo domains) by site-directed mutagenesis. Analysis of the putative TRDs of the mono-specific M.*Hpa*II, M.*Msp*I and M.*Bsu*FI (all of which recognize the sequence CCGG) (212) have revealed that the TRD of M.*Hpa*II is substantially different from that of M.*Msp*I and M.*Bsu*FI. This is an interesting observation as M.*Hpa*II methylates the inner cytosine whereas

M.*Msp*I and M.*Bsu*FI both methylate the outer cytosine. This observation suggests that the TRDs function is not just to direct what target to recognize but also to direct which cytosine to be methylated (203).

Unlike the 5^{m}C-methyltransferases, the $N4^{m}$C and $N6^{m}$A methyltransferases show much higher sequence diversity among themselves. Comparison of the amino acid sequences of $N4^{m}$C and $N6^{m}$A methylases have revealed only two conserved segments, namely F-X-G-X-G and (NSD)-P-P-(YF) (30, 100, 115). Conserved segment I (F-X-G-X-G) has been suggested to be involved in AdoMet binding because of its homology to a similar motif in 5^{m}C methylases and other AdoMet-binding proteins (85). Conserved segment II ((NSD)-P-P-(YF)) has been suggested to be involved in the transfer of methyl groups to the adenines or cytosines of the target DNA. Crystal structures of M.*Hha*I (34) and M.*Hae*III (155), both of which are 5^{m}C cytosine methylases, and of M.*Taq*I (108), a $N6^{m}$A methylase and M.*Pvu*II (129), a $N4^{m}$C methylase have been resolved recently. Comparisons of the crystal structures have indicated that motifs I to III and X are responsible for AdoMet binding (34, 102, 108, 168). Motifs IV, VI and VIII have been observed to be responsible for catalysis (168), as they form an active site along with motifs V and VII. Although the primary amino acid sequence comparisons failed to highlight conserved domains, structural comparisons of these methylases with catechol-O-methyltransferase (COMtase) and other DNA methylases have shown striking similarity, in-spite of the fact that these methylases methylate different substrates (144). Guided by this common domain structure, Malone, *et al.* (129) have compared 33 $N6^{m}$A, nine $N4^{m}$C, one 5^{m}C and one COMtase with each another. They were able to identify nine segments of sequence similarity corresponding to motifs I to VIII and X in

the $5^{m}C$ methylases. However, they were not able to identify a region homologous to motif IX of the $5^{m}C$ methylases. Based on the analysis they observed that the methylases differ in the relative linear order of the three regions, the AdoMet binding region (motifs I to III and X), the catalytic site (motifs IV to VIII) and the target recognition region. Thus, they have grouped the methylases into three groups, α, β, and γ based on the order in which these regions are present. In group α the AdoMet binding region (motifs X-I-II-III, in that order) is at the amino terminus followed by the target recognition region and then the catalytic region (motifs IV-V-VI-VII-VIII, in that order). In case of β, the order is catalytic region, target recognition region and then the AdoMet binding region. In case of γ, the AdoMet binding region is at the N-terminus followed by the catalytic region and then the target recognition region. The difference between $5^{m}C$ methylases and the methylases that belong to group γ is that in case of group γ motif X is at the N-termini whereas it is at the C-termini in the case of $5^{m}C$ methylases. This circularly permuted motif order is not very surprising for two reasons: (a) It had been pointed out, before the crystal structures were resolved, that the two previously identified conserved segments I (suggested to be involved in AdoMet binding) and II (suggested to be the catalytic site) appeared in different orders in various methylases (100, 216, 217). (b) It has been shown that there are some $5^{m}C$ methylases that have two separate subunits (each consisting of different motifs) and both the subunits have to be present for functionality (95, 122). It has also been shown that regions of the methyltranferase when expressed separately could complement each other if allowed to interact with each other *in vivo* (148).

CHAPTER III

MATERIALS AND METHODS

Bacterial Strains

The laboratory strain that is equivalent to the wild-type *S. pneumoniae* used in this study is a non-encapsulated strain Rx1 (97). Rx1 is derived from R36A strain of Avery (5). The pnuemococcal strains used in this study are listed in Table I.

S. pyogenes ATCC 19615 used for conjugation experiments in this study was provided kindly by the Stillwater Medical Center.

E. coli strains used in this study are listed in Table II.

Growth and Storage of Bacterial Strains

Streptococcal strains were grown in casein hydrolysate (CAT) medium (10g of casein hydrolysate, 5g tryptone, 5g NaCl, and 1g yeast extract in one liter of distilled water). The medium was autoclaved, cooled, and supplemented with sterile, 20% glucose and 0.5 M K_2HPO_4 to yield a final concentration of 11 mM and 16 mM, respectively. The supplemented medium was called CATPG broth or agar (1.5% w/v).

TABLE I

Streptococcal Strains

STRAIN	GENOTYPE	REFERENCE
S. pneumoniae		
Rx1	*hex*$^-$	175
DP1002	*nov*-1	63
DP1004	*str*-1	63
DP1322	Tn*5253* (*cat tet*)	181
DP1343	*nov*-1, Tn*5253* (*cat tet*)	181
DP1402	*str*-1, BM4200 (*cat tet erm aph*A)	181
DP1617	*hex*$^-$, *str*-1, *ery*-2, *nov*-1, *fus*, *sulf*-d	181
DP3111	*hex*$^+$, *str*-41, *Dpn*II$^+$	64
DP3218	*rif*, pIP501 (Cm Em)	181
SP1000	*str*-1, *fus*, Tn*5252* (*cat*)	7
SP1206	*nov*-1, Tn*5252* (*cat* Ω Em) Tra$^-$	This study
SP1207	*nov*-1, Tn*5252* (*cat* Ω *aph*A) Tra$^-$	This study
SP1208	*nov*-1, Tn*5252* (*cat* Ω *aph*A) Tra$^-$ pIP501 (Cm Em)	This study
SP1209	*nov*-1, Tn5252 (*cat* Ω *aph*A) Tra$^-$, pIP501 (Cm Em), pLS1 (*tet*)	This study

Table I continued

SP1210	*nov*-1, Tn*5252* (*cat* Ω *aph*A) Tra$^-$, pIP501 (Cm Em), pSJ017 (*tet*)	This study
SP1211	*nov*-1, Tn*5252* (*cat* Ω *aph*A) Tra$^-$, pIP501 (Cm Em), pSJ117 (*tet*)	This study
SP1212	*nov*-1, Tn*5252* (*cat* Ω *aph*A) Tra$^-$, pIP501 (Cm Em), pSJ119 (*tet*)	This study
SP1213	*str*-1, Tn*5252* (*cat* tΩ Em) Tra$^-$	This study
SP1214	*str*-1, Tn*5252* (*cat* Ω*aph*A) Tra$^-$	This study
SP1254	*str*-1, Tn*5252* (*cat* Ω Em) Tra$^+$	97
SP1255	*str*-1, Tn*5252* (*cat* Ω Em) Tra$^-$	97
SP1262	*str*-1, Tn*5252* (*cat* Ω Em) Tra$^+$	This study
SP1264	*nov*-1, Tn5252 (*cat*) Tra$^+$	This study
SP1265	*nov*-1, Tn*5252* (*cat* Ω Em) Tra$^+$	This study
SP1273	*nov*-1, Tn*5252* (*cat* Ω Em) Tra$^+$	This study
SP1292	*str*-1, Tn*5252* (*cat* Ω Em), pLS1 (*tet*) Tra$^+$	This study
SP1293	*str*-1, Tn*5252* (*cat* Ω Em), Tra$^+$ pSJ107 (*tet*)	This study
SP1294	*str*-1, Tn*5252* (*cat* Ω Em), Tra$^+$ pSJ117 (*tet*)	This study
SP1295	*str*-1, Tn*5252* (*cat* Ω Em), Tra$^+$ pSJ119 (*tet*)	This study

Table I continued

SP1296	*nov*-1, Tn*5252* (*cat* Ω Em), Tra^+ pLS1 (*tet*)	This study
SP1297	*nov*-1, Tn*5252* (*cat* Ω Em), Tra^+ pSJ107 (*tet*)	This study
SP1298	*nov*-1, Tn*5252* (*cat* Ω Em), Tra^+ pSJ117 (*tet*)	This study
SP1299	*nov*-1, Tn*5252* (*cat* Ω Em), pSJ119 (*tet*)	This study
SP1322	*nov-1*, pSJ107 (*tet*)	This study
S. pyogenes		
ATCC 19615	*opt*	Stillwater Medical Center

nov-1	Chromosomal point mutation conferring resistance to novobiocin.
fus	Chromosomal point mutation conferring resistance to fusidic acid.
str-1, *str*- 41	Chromosomal point mutation conferring resistance to streptomycin
ery-2	Chromosomal point mutation conferring resistance to erythromycin
sulf-d	Chromosomal point mutation conferring resistance to sulfonamide
*aph*A	Aminoglycoside phosphoryl transferase type A (resistance to kanamycin)
cat	Chloramphenicol acetyl tranferase (resistance to chloramphenicol)
tet	Resistance to tetracycline (type M)
Em	Resistance to erythromycin
opt	Resistance to optochin

TABLE II

***E.coli* Strains**

STRAIN	RELEVANT GENOTYPE
JM109	F', *rec*A1, *lacZ*^{q} Δ (*lacZ*)M15, McrA^{-}, r_K^{-} m_K^{+}
C600	F^{-}, *lac*Y1, McrA^{-}, r_K^{+} m_K^{+}, McrBC^{+}
GM2163	F^{-}, *lac*Y1, *dam*13::Tn*9*, *dcm*-6, *mcr*B1, *hsd*R2 (r_K^{-} m_K^{+}), *mcr*A
ER2508	*lon*:: Tn*10*, Δ(*mal*B), Δ(*mcr-mrr*)20
XL1-Blue MRA	Δ(*mcr*A)183, Δ(*mcr*CB-*hsd*SMR-*mrr*)173
XL1-Blue MRF' Kan	Δ(*mcr*A)183, Δ(*mcr*CB-*hsd*SMR-*mrr*)173, *rec*A1, *lac*[F' *pro*AB *lac*I^{q}ZΔM15 Tn*5* (Kan^{r})]
BL21 (DE3)	F^{-} ,*omp*T, *hsd*SB(r_B^{-} m_B^{-}), *gal*, *dcm* (DE3)

S. pnuemoniae cultures were grown to an OD_{550nm} = 0.2 (which is approximately 2.0×10^8 CFU/ml). Due to the induction of autolysis at the stationary phase, pneumococcal cultures to be stored were never allowed to grow beyond OD_{550nm}= 0.2. *S. pyogenes* was grown overnight prior to storage. The cultures were stored at –80°C with 10% (v/v) sterile glycerol.

E. coli strains were grown in Luria-Bertani (LB) medium (10g tryptone, 5g yeast extract, and 10g NaCl in one liter of distilled water). The pH of the medium was adjusted to 7.5 with NaOH and the medium sterilized. The agar concentration used was 2.0% w/v. For short-term storage, overnight cultures of *E. coli* were stored at -20°C with 50% v/v glycerol while long-term storage was at -80°C with 10% glycerol. All the bacterial cultures were grown at 37°C unless and otherwise indicated. Each of the above mentioned media was supplemented with the appropriate antibiotic, whenever there was a need to do so. The concentrations of the antibiotics used to select bacterial strains used in this study are tabulated in Table III.

Plasmids

Vector plasmids used in this study were pBluescript SK^+, pVA891, pLS1, pMAL-2 and pET-30^+ a-c. pVA891 (126) is an *E.coli* plasmid carrying Cm^r and Em^r. When inserted into the chromosome of pneumococcus, it expresses Em^r. The broad-host-range plasmid pLS1 (190) is a derivative of pMV158 isolated from *Streptococcus agalactiae* (25). pLS1 can autonomously replicate and express tetracycline resistance in

Table III

Selective Antibiotic Concentrations

Genotype/ Phenotype	Antibiotic	Concentration (μg/ml) Stab Plate/Broth	Overlay
Streptococcus pneumoniae			
cat (Tn*5252*)	chloramphenicol	5	15
Em^r	erythromycin	3	5
nov	novobiocin	10	10
str	streptomycin	200	200
rif	rifampicin	10	10
*aph*A (Tn*1545*)	kanamycin	200	1000
tet	tetracycline	2	5
Cm^r (pIP501)	chloramphenicol	3	5
Streptococcus pyogenes			
cat (Tn*5252*)	chloramphenicol	5	
Em^r	erythromycin	5	
opt	optochin	20	
Escherichia coli			
Cm^r	chloramphenicol	15	

Table III continued

Cm^r (pLysS)	chloramphenicol	34
Em^r	erythromycin	200
Tc^r	tetracycline	10
Kan^r (pJH1)	kanamycin	50
Kan^r (Tn*1545*)	kanamycin	25
Kan^r (pET30)	kanamycin	70
Amp^r	ampicillin	100
Sm^r	spectinomycin	300

S. pneumoniae (167) and in recombination-proficient *E. coli* (109). The prokaryotic expression vector plasmids pMAL-2 and pET-30^{+} a-c plasmids were purchased from New England Biolab (NEB) and Novagen, respectively. The relevant features of these and other plasmids used are listed in Table IV.

Enzymes, Chemicals and Antibodies

Restriction endonucleases, DNA molecular weight standards, and most of the commonly used modifying enzymes were purchased from one of the following suppliers: Bethesda Research Laboratories (BRL), Promega Biotech, New England Biolab (NEB), and United States Biochemical Company (USBC). Sequenase Version 2.0 T7 DNA Polymerase and Deep Vent DNA Polymerase were purchased from USBC and NEB respectively. DNase I, RNase , antibiotics, and bovine serum albumin (fraction V) (BSA) were purchased from Sigma Chemical Company. Media and agar for bacterial growth were purchased from Difco and USBC. Agarose for horizontal gel electrophoresis was purchased from Fisher. Molecular biology grade ultrapure agarose from Bio-rad Laboratories was used for preparatory gels. {α–^{32}P}-dCTP, {α-^{35}S}-dATP, {^{35}S}-Methionine, {^{3}H}-S-Adenosyl L- Methionine, and {^{14}C}- methylated proteins were purchased from New England Nuclear research Products (NEN). Primers for sequencing and PCR reactions were made at the Recombinant DNA/Protein Resource Facility at Oklahoma State University.

Table IV

Plasmids and Cloning Vectors

PLASMID	RELEVANT FEATURE	ORIGIN/REFERENCE
pBluescript SK^+	Amp^r, lacZα	Stratagene
pVA891	Cm^r, Em^r	126
pLS1	Tc^r	190
pMal-2	Amp^r, lacZα	New England Biolabs
pET30 a-c^+	Kan^r	Novagen
pRL425	Cm^r, Em^r, Amp^r	56
pDG792	Kan^r (from pJH1), Amp^r	ATCC
pDL276	Kan^r (from pJH1), lacZα	Dr. J.J. Ferretti
pLysS	Cm^r, T7 $lysozyme^+$	Novagen
pAT21-1	Kan^r (from Tn*1545*)	Dr. Patrice Courvalin
pIP501	Cm^r, Em^r, Tra^+	64
pSJ107	[pLS1 *Eco*RI::3.27 *Eco*RI], Tc^r	This study
pSJ114	[SK^+ *Eco*RI/*Kpn*I:: 1.05 *Eco*RI/*Kpn*I], Amp^r	This study
pSJ115	Exonuclease/S1 derivative of pSJ114, Amp^r	This study

Table IV continued

pSJ116	Exonuclease/S1 derivative of pSJ114, Amp^r	This study
pSJ117	pSJ107 with the unique *Kpn*I site Klenowed, Tc^r	This study
pSJ119	pSJ107Δ1.8 kb *Xba*I fragment, Tc^r	This study
pSJ126	[pRL425*Hind*III ::*Bss*HII/*Ava*I Kan^r of pDL276], Amp^r, Kan^r	This study
pSJ133	[pMalC$_2$ *Xmn*I:: ORF6α] (reverse orientation), Amp^r	This study
pSJ136	[pMalP$_2$ *Xmn*I:: ORF6α], Amp^r	This study
pSJ137	[pET30b*Eco*RV/*Hind*III::pSJ136 1.4kb*Ecl*136II/*Hind*III], Kan^r	This study
pSJ144	[pSJ119*Xba*I::pVA891*Xba*I], Tc^r, Cm^r, Em^r	This study
pSJ150	[pDG792*Cla*I Amp^r::pAT21-1 *Cla*I Kan^r], Amp^r, Kan^r	This study
pSJ170	[pSJ126*Hinc*II Amp^r:: pSJ150 *Nru*I/*Sma*I Kan^r], Amp^r, Kan^r	This study

Table IV continued

pSJ172	[pSJ119*Xba*I Tc^r::pSJ170*Xba*I Kan^r], Tc^r, Kan^r	This study

Amp^r	Resistance to ampicillin
Cm^r	Resistance to chloramphenicol
Kan^r	Resistance to kanamycin
Tc^r	Resistance to tetracycline
Em^r	Resistance to erythromycin
*lacZ*α	Capable of α-complementation enabling chromogenic differentiation
T7 lysozyme$^+$	The plasmid pLysS has the gene for making this enzyme

Polyclonal antibodies against ORF6α (generated by injecting purified protein into mice) were obtained from the Hybridoma Center at Oklahoma State University. The anti-mouse secondary antibody conjugated with alkaline phosphatase was a generous gift by Dr. Kay Scheets. Miscellaneous chemicals and reagents were purchased from Fisher, Sigma USBC, and VWR.

Transformation

S. pneumoniae Transformation

S. pneumoniae cells were made competent according to Guild and Shoemaker (63). The competent cells were stored at -80°C in 10% glycerol. For transformation, the cells were thawed in an ice-water bath and mixed with transforming DNA (1 μg/ml for chromosomal DNA and 10 μg/ml for plasmid DNA). Transformation was carried out at 37°C for 30 min. Transformation was stopped by adding DNase I to a final concentration of 25 μg/ml and incubating the mixture for five min at 37°C. Appropriate dilutions of the transformed cells were made in CATPG broth. Appropriate volumes were then added to 4 ml CATPG broth with 2% bovine blood or sheep blood and mixed with 4 ml of CATPG agar cooled to 49°C (Cell layer), and poured on a 20 ml base CATPG agar. After the cell layer solidified, 10 ml of CATPG agar was overlaid (Buffer layer) and incubated for 90 min at 37°C to allow phenotypic expression (208). The plates were then overlaid with 10 ml CATPG agar containing the appropriate antibiotic and incubated overnight at the same temperature.

Competence in *S. pneumoniae* is due to the release of competence factors into the surrounding media. Pneumococcal cells can be made competent by the addition of the supernatant of competent cells as they contain these competence factors. As preparation of pneumococcal competent cells is a tedious procedure, in some cases, competent cells were not prepared but the supernatant of Rx1 (the wild-type strain) competent cells was used as described. Host cells to be transformed were grown in CATPG media to an OD_{550nm} = 0.2 and 100 µl of the culture was centrifuged (14,000 rpm, 5 min) and kept on an ice-water bath. Rx1 competent cells were centrifuged (14,000 rpm, 5 min) and the supernatant was transferred to a fresh microfuge tube at 0°C. This process was repeated once more and then 500 µl of the competent Rx1 supernatant was added to the pellet of the host cells at 0°C and mixed. The transforming DNA was added and the mixture was incubated at 37°C for 2 h and then plated using the overlay technique.

E. coli Transformation

Competent *E.coli* cells were prepared and transformed according to the method described by Hanahan (71). The competent cells were stored for two months at -80°C. Appropriate volumes of transformed cells were plated immediately on LB agar plates containing the selective antibiotic. One hundred microliter of 20 mg/ml 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside (X-gal), and 10 µl of 20 mg/ml of isopropylthio-β-D-galactoside (IPTG) were used for chromogenic differentiation of recombinants and these were spread on the LB agar surface (100 mm) and dried before use.

Conjugation

Conjugation between pneumococcal donors and recipients was performed by the method described by Smith and Guild (183) with minor modifications. The cells were grown to an OD_{550nm} = 0.2 in Difco CATPG broth supplemented with 0.001 % choline chloride. The donors and recipients were mixed at ratio of 1: 5 in the presence of 10 mM $MgSO_4$, 2 mg/ml BSA, and 100 µg of DNase I. Approximately 1x 10^9 cells were filtered onto a nitrocellulose filter (Millipore: 13 mm diameter and 45µm pore size). The filters were placed cell side down on Difco CATPG agar containing 10 mM $MgSO_4$, 2 mg/ml BSA, 1 mM $CaCl_2$, and 100 µg/ml DNase I, and overlaid with the same agar. Each mating was done in triplicate and repeated at least one more time to check for reproducibility. Incubations were carried out at 37°C for 4 h maximum. The cells were harvested by cutting out both layers of agar with the filter. Each filter was then resuspended in 3 ml of resuspension broth (CATPG containing 10 mM $MgSO_4$, 2 mg/ml BSA, 100 µg/ml DNase I and 10% Glycerol). After vortexing thoroughly, the liquid portion was transferred to a tube and stored at -80°C for later plating. Appropriate dilutions of the conjugation mix were plated by the overlay method with selection and scored for the transconjugants as well as the parental strains. While selecting for Kan^r transconjugants, the 90 min expression time was given for phenotypic expression but without the buffer layer. After 90 min, the drug layer containing appropriate antibiotic(s) was poured over the cell layer.

Conjugations between pneumococcal strains and *S. pyogenes* strains were essentially performed the same way. To screen for the transconjugants (*S. pyogenes* ΩTn*5252*) and the recipient (*S.pyogenes*), appropriate volumes of the mating mixture were spread on the surface of the plate. The donor *S. pneumoniae* was plated by the overlay method of plating.

The transconjugants were individually checked for unselected markers by replica plating on selective CATPG agar with 2% bovine blood or sheep blood. If less than 50 transconjugants were observed, all of them were replica plated, if more, then 50 were selected at random for replica plating.

Chromosomal and Plasmid DNA isolation

Chromosomal DNA isolation from *S. pneumoniae*

To isolate the chromosomal DNA, *S. pneumoniae* cells were grown in CATPG to an OD_{550nm} of 0.4, and ethylenediamine tetraacetic acid (EDTA), was added to a final concentration of 10 mM. After a ten minute incubation at 0°C, the cells were centrifuged at 5,000 rpm in a Sorvall GSA rotor at 4°C for 10 min. The pellet was washed once in 25 ml of 1x SSC (0.15 M sodium chloride, 15 mM sodium citrate) and centrifuged at 5,000 rpm for 10 min. The pellet was resuspended in 4 ml of 2X SSC and 5 ml of solution I {5 mM EDTA, 0.1% Sarkosyl, 10 mM Tris-HCl (pH 8.0), 1% Triton-X-100 and 100 mg/ml RNase} was added and incubated at 37°C for 10 min or until the solution cleared. 1 ml of lysis mix {0.2% sodium deoxycholate (DOC), 0.1% Sodium

dodecyl sulfate (SDS) in 1x SSC} was added and incubated at 65°C for 30 min. The lysate was purified according to the method described by Marmur, *et al.* (132) with some modifications. Briefly, the viscous lysate was extracted twice with an equal volume of phenol: chloroform-isoamyl alcohol (25:24:1), twice with chloroform-isoamyl alcohol (24:1), and precipitated with two volumes of 95% ethanol at -20°C for 30 min. The DNA was pelleted by centrifugation at 4°C at 20,800xg for 10 min. The pellet was vacuum dried at 65°C and resuspended in TE buffer (10 mM Tris-HCl (pH 8.0), 1 mM EDTA). The concentration of the DNA was measured in a UV spectrophotometer (Spectronic 1001, Milton Roy Co.) and stored at -20°C.

Plasmid DNA isolation from *S. pneumoniae*

For small plasmid preparations, single colony isolates of *S. pneumoniae* were grown in 6 ml CATPG broth with appropriate selection to an OD_{550nm} of 0.4. A 3-ml aliquot of the culture was centrifuged at 11,750 x g for one minute and the supernatant discarded. The pellet was resuspended in 50 µl 1x SSC. Fifty microliters SDS/DOC [(0.2% w/v) SDS and 0.5% (w/v) DOC] were added to the cell suspension and incubated at 37°C until lysis was observed. The lysate was boiled for five min and 10 µl of the sample was electrophoresed on an agarose gel.

Large-scale plasmid purification was done by growing a 250 ml culture of *S. pneumoniae* (supplemented with 0.001% choline chloride) to an $OD_{550nm} = 0.4$. Ten milliliter of 0.5 M EDTA was added to the culture at 0°C and the cells were centrifuged at 7,000 x g in a Sorvall GSA rotor for 10 min at 4°C. The cells were washed once in

30 ml 1X SSC and pelleted before resuspension in 1 ml 0.1X SSC. Five and one-half milliliter of lysing solution containing 1 volume 10X lysis buffer [1% (w/v) DOC, 0.1% (w/v) SDS], 1 volume of 1M Tris-HCl (pH 8.0) and 9 volumes 10X SSC were added and the cells were incubated at 37°C until lysis. Seven and one-half milliliter of 4.4% (w/v) SDS solution [4 volumes of 20% (w/v) SDS, 5 volumes of 10X SSC, and 1 volume of Tris-HCl, pH 8.0] was added to the lysate and incubation was continued at 65°C until the lysate cleared. To the lysate, 5.0 ml of 4M NaCl was added, and the lysate was left overnight at 4°C. The following day, the lysate was centrifuged for 40 min at 4°C at 11,000 x g. 100 µl of RNase (10 mg/ml) was added to the supernatant and the lysate was extracted once with chloroform-isoamyl alcohol (24:1). The plasmid DNA in the aqueous phase was precipitated at -20°C for an hour with 95% v/v ethanol, centrifuged and resuspended in 4 ml TE buffer (10 mM Tris-HCl, pH 8.0, 1 mM EDTA). Cesium chloride (1.05 g/ml) and 400 µl of ethidium bromide (10 mg/ml) were added and centrifuged at 227,640 x g in a VTi65.2 rotor at 17°C for 18 h in an XL-70 ultracentrifuge (Beckman). The plasmid DNA band was extracted using a 22-G needle. To remove the ethidium bromide, the fraction was further extracted with water saturated butanol, and dialyzed overnight in 4 L of TE buffer. The concentration of the DNA was measured in a UV spectrophotometer (Spectronic 1001, Milton Roy Co.) and stored at -20°C.

Plasmid DNA isolation from *E. coli*

Small-scale plasmid DNA isolation was done as described by Sambrook, *et al.* (163). A single colony of *E. coli* was inoculated in 3 ml (or 5 ml) of LB with appropriate antibiotic selection and incubated overnight at 37°C. One and one-half milliliter (or 3 ml) of the overnight culture was centrifuged at 11,750 x g at room temperature for 30 secs. Cells were resuspended in 200 µl of cell resuspension solution [50 mM glucose, 25 mM Tris-HCl (pH 8.0), and 10 mM EDTA] by vortexing vigorously and kept at room temperature for 5 min. Cells were lysed with 200 µl of freshly prepared lysis solution [0.2N NaOH, 1% SDS] and kept at room temperature for another 5 min. The lysate was mixed with 200 µl of neutralization solution [3 M potassium acetate (pH 5.2)] and held on ice for five min. The lysate was centrifuged at 11,750 x g at room temperature, twice for 5 min. The supernatant fluid was transferred to a fresh tube. Plasmid DNA was precipitated with two volumes of 95% ethanol at -20°C for 15 min. The DNA was pelleted, vacuum dried and resuspended in 50 µl TE buffer. Ten microliters of the plasmid DNA was electrophoresed on an agarose gel to check for recombinant plasmids. For restriction endonuclease analysis, the plasmid DNA was extracted using phenol-chloroform prior to ethanol precipitation.

Large-scale plasmid purification was performed essentially in the same manner as the miniprep except for the use of a 250 ml culture for scale-up and a CsCl density gradient to separate the plasmid from the genomic DNA. In case of low copy number plasmids, an additional plasmid amplification step was performed by growing the cells for a further 12 h in the presence of spectinomycin (300 µg/ml). Cesium chloride at

1.0 g/ml and ethidium- bromide at 1 mg/ml were added to the DNA resuspended in TE buffer, centrifuged at 227,640 x g in a VTi65.2 rotor at 17°C for 18 h (or at 383,700 x g for 4 h) in the XL-70 ultracentrifuge (Beckman). The plasmid DNA band was extracted using a 22-G needle, to remove the ethidium bromide it was further extracted with water saturated butanol, and dialyzed overnight in TE buffer.

DNA Manipulation

Cloning

The DNA fragments of Tn*5252* were subcloned in plasmid vectors and transformed into *E.coli* host cells, and in some cases, competent *S. pneumoniae* cells. Ligations were carried out in 20 µl volume with less than 5 µg DNA at a ratio of one vector to five inserts at 16°C for 24 h. Ligations that were used for transforming competent pneumococcal cells were carried out in a reaction volume of 60 µl with approximately 10 µg of DNA with a vector:insert ratio of 1:5. For insertion mutagenesis, the insert:vector ratio was 1: 1 in a 60 µl volume. After overnight ligations, the ligation mixture was extracted with phenol-chloroform, linearized with restriction enzyme and used as donor DNA in transformation of competent pneumococcal cells.

Agarose Gel Electrophoresis

DNA samples were analyzed at room temperature by separation on horizontal agarose gels using 1/2X TBE (45 mM Tris-borate, and 1 mM EDTA, pH 8.3). Depending on the size of the DNA, the concentration of the agarose varied from 0.8%-1.2%. The DNA samples were mixed with tracking dye containing 5% (v/v) glycerol, 3 mM EDTA (pH 8.0), bromophenol blue (0.04%) and xylene cyanol (0.04%), before loading. Typically, the gels were run at 5 volts/cm constant electric current for approximately 90 min. After electrophoresis, the separated DNA samples were observed on a 300 nm UV illuminator after the gels were stained in ethidium- bromide (1 μg/ml final concentration) for 10 min followed by destaining in deionized water for the same period. Photographs were taken using a Kodak Polaroid camera or an Alpha Imager 2000 of Alpha Innotech Corporation. Molecular weight standards (*Hin*dIII fragments of phage Lambda and *Hae*III fragments of phage ϕX174 or 1 kb ladder of BRL) were run along with the DNA samples for size comparison. Preparative minigels were used to purify specific DNA fragments, and 0.8% gels were made with ultrapure agarose (Bio-Rad). After staining and destaining, a gel slice containing the DNA fragment was electroeluted and purified according to Sambrook, *et al.* (163).

Conversion of 3' overhang to blunt-end with Klenow fragment

The single stranded 3' overhangs were converted to a blunt end using the 3'-->5' exonuclease activity of Klenow DNA Polymerase. The reaction mix consisted of 4.5 μg

of digested DNA, 1X Klenow buffer and Klenow DNA polymerase (1-5 units) in a final volume of 50 µl. The reaction was incubated at room temperature for 15 min. Following phenol:chloroform-isoamyl alcohol extraction, the DNA was precipitated with 1/2 volume 7.5 M ammonium acetate and 2 volumes ethanol for 20 min at -20°C and used in ligation reactions.

DNA Sequencing with T7 DNA Polymerase

(Sequenase version 2.0)

Dideoxy-chain-termination method of sequencing (163) was carried out with Sequenase version 2.0 (USBC) T7 DNA polymerase, and [α-^{35}S]dATP (12.5 mCi/ml) from NEN. M13F, M13R, pMalE were the standard primers used. Primers specfic for the transposon DNA under study were synthesized at the DNA/Protein Recombinant DNA/Protein Resource facility of Oklahoma State University.

Denaturing Polyacrylamide Sequencing Gel

Seven percent (w/v) polyacrylamide gel containing a concentration gradient in TBE buffer was prepared according to the methods described by Biggin, *et al.* (18). For electrophoresis, a Sequi-Gen Sequencing Cell from Bio-Rad was used. Both the glass plates were thoroughly cleaned and air dried. The gel plate was coated with "Glue" (10ml 95% ethanol, 30 µl glacial acetic acid, 50 µl of G-methacryloxy propyl trimethoxysilane and 0.5 ml of ddH_2O), and the buffer chamber plate was coated with

Sigmacote (Sigma). Gel solution A contained 17.5% (v/v) of 40% acrylamide solution (38% acrylamide and 2% bis-acrylamide), 25% (w/v) urea and 10% (v/v) TBE buffer. Immediately before pouring the bottom plug as well as the gel, 0.3% (w/v) ammonium persulfate (APS) and 0.05% TEMED (N, N, N',N'-tetramethylethylenediamine) were added to the solution. The plates were assembled and the bottom was plugged with the solution B [Solution A supplemented with 25% (w/v) of sucrose and 0.005% (w/v) of bromophenol blue]. The gel apparatus was filled with the denser solution B at the bottom of the gel at a ratio of 1:5 (solution B: solution A) and the chamber was filled with solution A. The gel was allowed to solidify at room temperature for at least 18 h.

Constructions of nested deletions for sequencing

Nested deletions were created of DNA fragments over 1.0 kb in size cloned in pBluescript vector. The recombinant plasmid was doubly digested with two restriction enzymes in such a way that they left a 5'- recessed end near the primer binding site and a 3' - recessed end near the end of the insert DNA. This was done to provide a substrate for the enzyme exonuclease III used for creating deletions as it contains a specific 3'---->5' exonuclease activity. The reaction mixture, a total volume of 100 µl, containing approximately 5 µg of double digested DNA, 10 µl of 10X exonuclease buffer (0.66 M Tris-HCl, pH 8.0 and 6.6 mM $MgCl_2$) and exonuclease III (450 U/reaction) on ice. The reaction was started by transferring the tube to 30°C water bath and 10 µl samples were removed every minute for 10 min and transferred to 0°C. The first five samples were

combined in one tube and the last five samples in another before adding 150 µl of S1 nuclease solution (150 U S1 nuclease, 1% glycerol, 60 mM NaCl, 1.3 mM $ZnSO_4$, and 8 mM potassium acetate, pH 4.6). The tubes were transferred to room temperature for 20 min. This was done to generate blunts ends as S1 nuclease acts on single stranded DNA. The reaction was terminated with 20µl of S1 stop buffer (50 mM EDTA and 0.36 M Tris base). The samples were extracted with phenol: chloroform-isoamyl alcohol and chloroform-isoamyl alcohol followed by ethanol precipitation. The DNA was ligated and transformed into *E.coli* JM109 cells. The deletion derivative plasmids were isolated, their sizes confirmed by restriction analysis and used for sequencing.

Preparation of Plasmid DNA templates

Double-stranded-DNA templates were prepared with the Wizard miniprep DNA purification system (Promega) according to the manufacturer's recommendations with some modifications. 10 ml of an overnight culture was centrifuged and resuspended in 200 µl of cell resuspension solution (50 mM glucose, 25 mM Tris-HCl (pH 8.0), and 10 mM EDTA). The cells were lysed by the addition of 200µl lysis solution (1% SDS and 0.2 N NaOH) at room temperature for five min. The lysate was next neutralized with 200 µl of 3.0 M potassium acetate (pH 5.2) for 5 min at room temperature. After centrifugation at top speed in a microfuge (13,000 x g), the supernatant was transferred to a fresh microfuge tube and 1 ml of DNA purification resin was added to it and mixed by inversion. The mixture was passed through a purification column using a 3 ml

disposable syringe. The column was washed once with 2 ml of column wash solution (80 mM potassium acetate, 8.4 mM Tris-HCl (pH 7.5), 40 μM EDTA, and 55% ethanol). The excess wash solution was removed by brief centrifugation and the DNA was eluted with 100 μl TE buffer and stored at -20°C. This protocol provided DNA that could be used for fluorescent sequencing at the Oklahoma State University Recombinant DNA/Protein Resource facility.

Denaturing the Plasmid DNA

The double-stranded-DNA (dsDNA) (1μg) was denatured with 0.2 N NaOH and 0.2 mM EDTA at room temperature. It was neutralized with 0.3 M Tris-HCl, pH 4.5, and 0.8 M sodium acetate on ice. The denatured DNA was precipitated with ethanol, pelleted, vacuum dried and stored at -20°C.

Priming of Denatured DNA

Annealing of template to primer was carried out by setting up a 10 μl reaction volume containing 1 μg DNA, 30 ng primer, and 1X Sequenase reaction buffer and incubating it at 37°C for 20 min. A number of synthetic primers were constructed at the Recombinant DNA/Protein Facility, Oklahoma State University. M13 forward and reverse primers (Promega), and malE primer (NEB) were also used.

Labeling of the Plasmid DNA

The primed DNA was transferred to 0°C and 3.5 µl of labeling mix and 2.0 µl of Sequenase version 2.0 were added. Labeling was allowed to proceed for 10 min at 0°C. The labeling mix was prepared by mixing 10 µl of [α^{35}-S] dATP, 16 µl of 0.1 M DTT and 34 µl of dGTP mix for obtaining long stretches of DNA sequences. For obtaining sequences closer to the primer binding site, 10 µl of dGTP mix, 10 µl of {α^{35}-S} dATP, 28 µl of 0.1 M DTT and 52.0 µl of double distilled water (ddH_2O) were mixed. The enzyme was diluted 1:7 with 1 µl of enzyme and 7 µl of enzyme dilution buffer.

Termination

Four tubes, labeled G, A, T, C were placed at 49°C in a warming block and to each tube 2.5 µl of the corresponding termination mix (ddNTP) was added. The labeled DNA (3.5 µl) was added to each of the four tubes, centrifuged, and incubated for 4 min at 49°C. All the tubes were transferred to room temperature and 4 µl of stop buffer (95% formamide, 20 mM EDTA, 0.05% bromophenol blue and 0.05% xylene cyanol FF) was added to each tube and stored at -20°C until ready for sequencing.

Denaturing Polyacrylamide Sequencing Gel Electrophoresis

Following polymerization, the gel was prerun for 1 hour or until the temperature reached 40-45°C, using 0.5X TBE in the chamber and 1X TBE in the bottom tray at 35 Watts constant power. Before loading, samples were denatured in a 95°C water bath for 2 min and immediately loaded on the gel. To get effective resolution of the sequences, 3 M sodium acetate was added to the bottom of the tray to yield a final concentration of 1 M after the bromophenol blue in the sample ran off. At the completion of electrophoresis, the gel apparatus was dismantled and the glass plate with the gel was fixed for an hour at room temperature in 10% glacial acetic acid. The gel was briefly rinsed with distilled water and dried overnight at 65°C. The gel was exposed to X-ray film (Kodak XAR-5 or BIOMAX MR) until sufficient exposure was obtained. The sequences were manually read by placing the autoradioGram on an illuminator.

Nucleotide Sequence Analysis and Molecular Modeling

To analyze DNA sequences for restriction sites, identification of open reading frames and predictions of amino acid sequences, MacVector 4.5.3 software proGram was used. Similarity searches and bestfit analyses of the deduced amino acid sequences of potential open reading frames were performed by using the UWGCG Sequence Analysis Software Package (Genetic Computer Group, Inc. Wisconsin) via the University of Oklahoma Computer Center or the internet.

To predict the structure of ORF6α, SWISS-MODEL (http://www.expasy.ch/spdbv/mainpage.htm) was used. Multiple sequence alignment was also done using this service.

Southern Hybridization

Blotting

DNA-DNA hybridizations were essentially performed according to the method described by Southern (186). DNA samples for blotting were first digested using the appropriate restriction enzymes. The fragments were separated on 0.8% agarose gels, as described earlier. The DNA fragments were denatured with 0.5 M NaOH for 30 min at room temperature. The DNA was transferred onto a Genescreen Plus nylon membrane (NEN) using a Model 785 Vacuum Blotter (Bio-Rad). After transfer, the membrane was soaked in 2X SSC for 5 min, air-dried, and stored in a dessicator under vacuum.

Pre-hybridization

Before hybridization with a radiolabeled probe, the membrane with the DNA was placed in a "seal-a-meal" bag to which 10 ml of the prehybridization solution (1g Dextran Sulfate, 0.58g NaCl, 1 ml 10% SDS and 8 ml ddH_2O) was added. The solution was spread evenly on the membrane, and after squeezing out air bubbles, the bag was

sealed and prehybridization was carried out in a 65°C water bath overnight with gentle shaking.

Nick translation

Labeling of DNA to be used as probe for Southern hybridization was carried out as follows. Probe DNA (0.5-1.0 μg) was nick translated in the presence of 0.5 μl of DNase I (0.5 μg/ml), 5 μl of 10X buffered dNTP mix (5 μM each of dATP, dTTP and dGTP), 5 μl {α-^{32}P} dCTP (10 mCi/ml), and 0.5 μl of *E. coli* DNA polymerase I (9000U/ml) in a 50 μl total volume. The DNA was labeled at 16°C for 2 h. The reaction was terminated by adding 25 μl of 0.5 M EDTA, 25 μl of sheared salmon sperm DNA (2 mg/ml), 50 μl of 7.5 M ammonium acetate and 50 μl of TE. The DNA was precipitated with 100% ethanol for 15 min at -20°C, pelleted, dried, and resuspended in 600 μl of TE. The efficiency of incorporation of the isotope was measured by diluting 1.0 μl of the labeled probe in 5 ml of liquid scintillation cocktail (Beckman) and measuring the radioactivity in an LS 6000 SC liquid scintillation counter (Beckman).

Hybridization

Labeled DNA (1-4 x 10^6 CPM) was mixed with 100 μl 20X SSC, 400 μl sheared salmon sperm DNA (2 mg/ml) and the volume was adjusted to 2.0 ml with ddH_2O in a screw capped tube. The probe was denatured by boiling for 10 min. The

denatured sample was cooled rapidly by swirling in an ice bath for 3 min. The plastic bag containing the prehybridized membrane was cut open and the denatured probe was added into the bag using a Pasteur pipette and resealed. Hybridization was carried out in a 65°C waterbath with gentle shaking for at least 18 h. Following hybridization the membrane was removed from the bag and washed twice with 2X SSPE buffer {17.53% (w/v) NaCl , 2.76% (w/v) of NaH_2PO_4, and 0.74% (w/v) of EDTA, pH 7.4} for 15 min at room temperature, twice with 2X SSPE, 2% SDS for 45 min at 65°C, twice in 0.1X SSPE buffer for 30 min at room temperature, and once in 3 mM Tris base (unbuffered) solution for 15 min at room temperature. The membrane was briefly air-dried, and exposed to X-ray film (Kodak X-OMAT AR) at -80°C for varying time periods. The exposed film was developed according to the manufacturer's recommendation.

In vitro Transcription and Translation

In vitro transcription and coupled translation were performed using the S30 extract system of Promega using circular DNA as the template.

E.coli S30 Extract system for circular DNA

The standard reaction consisted of 4 µg of DNA template, 5 µl amino acid mixture lacking methionine, 20 µl S30 premix (lacking aminoacids but containing all other required components, including NTPs, tRNAs, an ATP generating system, IPTG and appropriate salts), 1 µl {^{35}S}- methionine (1,200 Ci/mole at 15 mCi/ml), 15 µl S30

extract and nuclease-free water to bring the final volume to 50µl. After mixing the components by a quick spin, the reaction was allowed to proceed for 2 h at 37°C. The reaction was stopped by transferring the tubes to 0°C for 5 min.

pBESTluc DNA was included in the protocol as the positive control. A negative control with pLS1 alone, the vector used for the cloning of passenger DNA, was also included.

Gel Analysis of Translation Products

Once the S30 extract reaction was completed, a 5 µl aliquot was removed for acetone precipitation. 20 µl of acetone was added to the reaction mix and incubated at 0°C for 15 min. The pellet was dried for 15 min under vacuum, resuspended in 20 µl sample buffer, and heated at 100°C for 5 min before loading on a SDS-polyacrylamide gel.

Discontinuous SDS-Polyacrylamide Gel Electrophoresis (PAGE)

Discontinuous SDS-PAGE was performed and described by Laemmli (111). A stock solution of 22% (w/v) acrylamide A (BRL) was prepared by dissolving 100g acrylamide and 2.7g bis-acrylamide (Fisher) in a total volume of 500 ml of ddH_2O and stored at 4°C. The concentration of acrylamide in the separating gel varied from 10-12% (w/v) depending on the size of the proteins to be separated; however, the stacking gel was always 3% (w/v). To make the electrophoresis mixture for the separating gel, 10.0

ml of 1.5 M Tris-HCl (pH 8.8), 0.4 ml of 10% SDS, a calculated volume of acrylamide A to yield the required concentration and ddH_2O to make a final volume of 40 ml were mixed and deaerated for 15 min. Before pouring the mixture into the gel casting assembly, 0.4 ml of 5% ammonium persulfate (APS) and 40 µl of TEMED were added. The acrylamide was overlaid with water saturated butanol and allowed to polymerize for 30 min. The stacking gel was made by mixing 1.35 ml Acrylamide A, 0.1% SDS, 2.5 ml of 0.5 M Tris-HCl (pH 6.8) and 6 ml of ddH_2O. After deaeration, 150 µl APS and 15 µl TEMED were added and poured over separating gel. Electrophoresis was carried out in Tris-Glycine-SDS buffer, using Life Technologies Model 250 power supply, at 20 mA constant current through the stacking gel and then the current was increased to 45 mA when the blue dye has entered the separating gel. Typically, one run was completed in 3-4 h. For rapid screening of the protein products or western blots, the mini-protean II gel electrophoresis apparatus (Bio-Rad) was used which had a run time of less than an hour.

Samples for electrophoresis were prepared by either pelleting 1 ml of cells and resuspending in 100 µl 1X sample buffer or, if collecting fractions, 100 µl of the fraction was mixed with an equal volume of 2X sample buffer. Broad range (2-212 kDa) protein markers (NEB) were also run along with experimental samples as a reference for size determination in each gel. To visualize unlabeled proteins after electrophoresis, gels were stained overnight in 50% ethanol, 10% glacial acetic acid, and 0.1% Coomassie brilliant blue R. The gels were then destained in warm 7% glacial acetic acid.

Fluorography

Following electrophoresis, labeled protein bands were visualized by fluorography. Fluorography was carried out according to the method of Laskey (114). In brief, the gels were soaked for 30 min in DMSO (Dimethyl sulfoxide) and this was repeated three times. This was then followed by soaking in 22% 2,5-diphenyloxazole (PPO)/DMSO for 3 h. The gels were finally washed in double distilled water for 1 h and vacuum dried on a 3 mm Whatman filter paper. The dried gel was exposed to X-ray film (Kodak X-OMAT AR) at -80°C and the film was developed following the recommendations of the supplier.

Western Blot

The electrophoretic transfer of the proteins from the SDS-PAGE gels was done using the mini Trans-Blot Electrophoretic transfer cell of Bio-Rad according to the manufacturer's recommendations. Briefly, transfer buffer (25 mM Tris, 192 mM Glycine) was made and stored at 4°C, at least overnight. After electrophoresis of the proteins, the discontinuous SDS-PAGE gel was removed from the glass plates and the stacking gel was trimmed off. The gel, cut nitrocellulose membrane, pre-cut filter papers and fiber pads were all soaked in the transfer buffer for 30 min at room temperature. The blot was assembled (as per the manufacturer's instructions) with the gel and the membrane sandwiched between filter papers and fiber pads and all the air bubbles were removed. The whole assembly was placed in the buffer tank and the tank was filled with

the transfer buffer. The frozen Bio-Ice cooling unit that was prepared earlier by freezing deionized, distilled water was also placed in the buffer chamber. A stir bar was also placed at the bottom of the buffer chamber for complete mixing. Then the power was turned on and run at 100 volts at 250 mA for one hour. After one hour, it was disassembled and the membrane was checked to make sure that the prestained protein markers had transferred. After this step, the membrane was rinsed in 50 ml of PBST (PBS with 0.05% Tween 20) for 15 min. At the end of 15 min, the solution was removed and appropriate amount of the antibodies (primary antibody) diluted in PBST (usually 1/200 of the final volume) was added and the blot shaken at room temperature for at least 1 h. The solution was discarded and the blot was washed 3 x with 50 ml PBST, five min each. The secondary antibody conjugated to alkaline phosphatase (rabbit antimouse IgG-alkaline phosphatase) diluted in PBST (1/5000) was added and the blot was shaken at room temperature for 2 h. The blot was washed 3 x in 50 ml PBST, 5 min each and washed once with PBS for 5 min. Then the blot was transferred to the alkaline phosphatase buffer [0.1M Tris-HCl (pH 9.5), 0.1M NaCl, 0.2 mM $MgCl_2$, 2 μM $ZnCl_2$ and 0.02% NaN_3] and left in that buffer until ready to expose to Lumiphos-530. On a clean, dry glass plate, 0.5 - 0.6 ml of Lumiphos-530 was spread (matching the width of the blot). The blot was taken from the alkaline phosphatase buffer, excess buffer was drained off, and placed carefully so that the protein side touched Lumiphos-530 and left for 1-2 min. The blot was removed, the excess of Lumiphos-530 was drained off and the blot was kept inside a plastic sheet protector and was exposed to an X-ray film (Kodak XAR-5 or BIOMAX MR) until the bands were visible. The blot was exposed to Lumiphos-530 again, if needed.

Preparation of *E.coli* total lysate

(for absorption of anti-*E.coli* antibodies)

The host cells used for purifying the ORF6α protein were BL21 (DE3) containing the plasmid pLysS, and the expression vector used for cloning ORF6α gene was pET30b. Total lysate from these cells was prepared to cross-react and remove the contaminating antibodies from the polyclonal antisera developed against ORF6α. The host cells containing the plasmids pET30b^{+} and pLysS were grown in 20 ml Rich broth at 37°C, overnight with appropriate selections. The next morning, 1 liter of fresh rich broth with 2.5 mM betaine HCl and appropriate selections was inoculated with 10 ml of the overnight grown culture and incubated at 37°C until an OD_{550nm} of 0.6 was reached. The cells were centrifuged at 11, 750 x g for 20 min and the pellet was resuspended in 100 ml of 1X SSC. The cells were centrifuged again for 20 min at 11, 750 x g and the pellet was resuspended in 10 ml solution of 20% sucrose and 50 mM Tris-HCl (pH 8.0). To this, 100 mg of lysozyme was added and incubated at 0°C for 10 min. 4 ml of 0.3 M EDTA was added and incubation was continued at 0°C for another 10 min. To this, 14 ml solution of 0.4% Triton-X-100, 30 mM EDTA, 50 mM Tris-HCl (pH 8.0), 50 μg/ml DNase I and 50 μg/ml RNase was added very slowly and swirled gently at room temperature for 15 min. 2 ml of 10% Triton-X-100 was added later and incubated for 20 min at 37°C. It was sonicated six times for 20 sec pulses and centrifuged at 23,000 x g for 30 min. The supernatant was stored in aliquots at -20°C.

Absorption of Anti-*E. coli* antibodies

The total lysate prepared was used for absorption as described in Sambrook, *et al.* (163). Briefly, the antiserum was diluted 1:10 with TNT [10 mM Tris-HCl (pH 8.0), 150 mM NaCl, 0.05% Tween 20] containing 1% gelatin and 5% nonfat dried milk. For every milliliter of this diluted antiserum, 0.5 ml of the lysate prepared was added and incubated at room temperature for 4 h. The absorbed antibody was stored at 4°C in the presence of 0.05% sodium azide until needed.

DNA Amplification with the Polymerase Chain Reaction (PCR)

DNA amplifications were performed with Deep Vent DNA polymerase (NEB). The basic PCR protocol was as recommended by the manufacturer with some modifications. PCR assays were done in 100 µl total volumes in 0.5 ml micro-centrifuge tubes. Transposon DNA from the streptococcal species under study served as a template DNA. A typical reaction mixture contained 10-100 ng of template DNA, two primers each at 0.2 - 1.0 mM concentration, 1X concentration of Vent reaction buffer {10 mM KCl, 2 mM $(NH_4)SO_4$, 20 mM Tris-HCl (pH 8.8), 2 mM $MgSO_4$ and 0.1% Triton X-100}, 1.25 mM each dNTP and 0.75 U Deep Vent DNA polymerase. The reaction mixture was overlaid with 100 µl sterile mineral oil, briefly centrifuged, and PCR performed in a Coy Tempcycler Model 60 (Coy Laboratory Products, Inc.).

On average, 30 cycles of amplification were carried out, each consisting of a 1 minute denaturation period at 94°C, 45 sec annealing at a temperature which was 5°C

lower than the calculated melting point of the primer with the lower Tm, and a variable extension period at 72°C. The Tm (also called Td; dissociation temperature), was calculated by the relation, Tm = 4(G+C) + 2(A+T), where the letters stand for each nucleotide in the primer. The extension period was determined by allowing 1 minute per kilobase of DNA at 72-75°C and, therefore, depended on the length of the DNA segment to be amplified. After the amplification 100 µl of TE buffer was added to the microfuge tube and briefly centrifuged, then approximately 195 µl of the bottom layer was transferred to a fresh microfuge tube and stored at -20°C until ready for gel purification.

5 µl of the amplified DNA was checked on an agarose gel to assess the quality and quantity of the amplified product. PCR amplified products were then purified by using preparative mini-gels (0.8%) made with ultra pure agarose (Bio-Rad). After destaining, a gel slice containing the DNA fragment was electroeluted into a dialysis bag and purified according to Sambrook, *et al.* (163).

Protein Purification

Protein purification using the pMAL P_2 system (pilot)

The strain containing plasmid pSJ136 was grown overnight in rich broth with ampicillin. From this overnight grown culture, fresh rich broth was inoculated with selection and grown to an OD_{600nm} of 0.5. One ml of this culture was centrifuged and the pellet was resuspended in 50 µl of 1X SDS-PAGE buffer and stored and this served as the uninduced control. To the rest of the culture, 0.3 mM IPTG was added and the

mixture incubated at room temperature for 3 h. A 500 µl sample was taken from this culture every hour until 3 h and was centrifuged. The pellet was resuspended in 100 µl of 1X SDS-PAGE buffer and stored, to compare the level of induction. The rest of theculture was split into two and was centrifuged. One of the pellets was resuspended in 4 ml of column buffer [20 mM Tris-HCl (pH 7.4), 200 mM NaCl and 1 mM EDTA], and the other was resuspended in 8 ml of a solution containing 30 mM Tris-HCl (pH 8.0) and 20% sucrose. The cells were stored at -20°C.

For obtaining the cellular fraction (crude extract), the cells stored in the column buffer were thawed and sonicated in short 15-sec pulses, 10 times with 1-minute intervals. The sonicated cells were centrifuged at 20,000 x g at 4°C for 25 min. The supernatant representing the cytoplasmic fraction was stored. The pellet obtained was resuspended in 4 ml 8 M urea and stored as the insoluble fraction. Five hundred microliter of the crude extract was mixed with the amylose resin and incubated on ice. After 15 min, it was briefly centrifuged, the supernatant was removed and stored, this served as the fraction that was not bound to the resin. The pellet was resuspended to 50 µl of 1X SDS-PAGE buffer and stored as the fraction bound to the resin. A small fraction of the resin bound crude extract (50 µl) was mixed with 10 mM Maltose and incubated for 10 min at room temperature. Then it was centrifuged and the supernatant was removed and stored. This was the fraction containing proteins released by maltose.

For obtaining the periplasmic fraction, 20 µl of 0.5 M EDTA was added to the cells stored in Tris/sucrose and incubated for 10 min at room temperature with shaking. Then the mixture was centrifuged at 11,000 x g at 4°C for 10 min. The supernatant was discarded, the pellet resuspended in 8 ml of ice-cold 5 mM $MgSO_4$, and shaken for

10 min in an ice-water bath. It was centrifuged at 4°C for 10 min at 11,000 x g. The supernatant is the cold osmotic shock fluid.

Purification of Target Protein Using the pET system

The protein of interest to this study was purified in larger quantities using the pET system using the BL21 (DE3) host containing the plasmid pLysS. The method used was as suggested by the manufacturer. Briefly, glycerol stock culture was used to streak a LB agar plate with appropriate antibiotics and incubated overnight. A single colony was picked from the plate, used to inoculate a 25 ml fresh rich broth with appropriate antibiotic selection and incubated with shaking until $OD_{600nm} = 0.6$. The culture was stored at 4°C overnight. The following day, 20 ml of the overnight culture was centrifuged and the pelleted cells were used to inoculate 1 liter of rich broth containing appropriate antibiotics and 2.5 mM betaine-HCl. The culture was incubated for 3 h at 37°C until the OD_{600nm} reached 0.6. A 1 ml sample was removed, the cells pelleted and resuspended in 100 µl of 1X sample buffer to serve as the uninduced control. The remaining culture was induced with IPTG at a final concentration of 1 mM and incubated at room temperature for 3 h. A 1 ml sample was collected, pelleted and resuspended in 100 µl of 1X sample buffer after every hour to compare the level of expression with that of an uninduced sample. At the end of 3 h, the flask was transferred to ice for 5 min and the cells were pelleted by centrifugation at 5000 x g for 5 min at 4°C. The pellet was resuspended in 30 ml 1X Binding buffer (5 mM imidazole, 0.5 M NaCl, 20 mM Tris-HCl pH 7.9) and 0.1% triton was added and stored at -20°C

overnight. After thawing at 0°C, 0.05% sarkosyl and 1 mM PMSF were added and the sample was sonicated for 4 min in 20 sec bursts with the tube in an ice-water bath to avoid heating of the sample that may lead to degradation of proteins. Sonication was continued until the sample was no longer viscous. The lysate was centrifuged for 30 min at 10,000 x g. 100 µl of the supernatant (soluble fraction) was mixed with 100 µl of 2X sample buffer and stored. A majority of the protein of interest was found to be in the insoluble fraction. Hence, the pellet was dissolved in 40 ml 1X binding buffer containing 6M urea. One hundred microliter of this insoluble fraction was mixed with 100 µl of 2X SDS sample buffer and stored. The urea-solubilized proteins were then filtered through a 0.45µ filter (Nalgene). The volume was made up to 100 ml with the same buffer and used to load the affinity column. One hundred microliter of this filtered insoluble fraction was mixed with 100 µl of 2X SDS sample buffer and stored.

To purify the protein from the soluble fraction, a two liter or a four liter culture was used, as the amount of induced target protein in the soluble fraction was very low. The protocol described above was followed to obtain the soluble fraction (200 ml) and this fraction was filtered before loading the affinity chromatography column. As before samples obtained during the purification process were stored for SDS-PAGE and methyltransferase assay.

Affinity Chromatography

To prepare the column, the resin (His. Bind metal chelation resin) was poured into a clean column and allowed to settle by gravity. The column was equilibrated with

seven column volumes of 1X charge buffer (50 mM $NiSO_4$). The column was equilibrated with three column volumes of 1X Binding buffer with 6M Urea. The prepared extract was loaded in the column and the column was washed with 10 volumes of 1X binding buffer with 6M urea. The flow rate was maintained at 10 column volumes per hour. One hundred microliter of the unbound sample was mixed with 100 µl of sample buffer to check for efficiency of binding. The bound protein was eluted using a gradient of imidazole, formed with 1X binding buffer with 6M urea and 1X elute buffer (1 M imidazole, 0.5 M NaCl, 20 mM Tris-HCl pH 7.9) with 6M urea. Ten microliter of each of the fraction was used to perform a Bradford assay to determine the concentration of purified protein in the fractions. 15 µl of each of the samples collected during the course of the process and the different fractions were loaded on a 12% SDS-PAGE gel to check the efficiency of expression and purification and to check for the presence of the protein.

The procedure for affinity chromatography of the soluble fraction was essentially the same as for the insoluble fraction except that the proteins were eluted with the elute buffer that did not contain urea. In some cases elute buffer containing 200mM, 400mM or 1M imidazole was used to elute the resin bound proteins.

Processing the purified sample

The fractions containing the protein peak were pooled and dialyzed to concentrate the protein as follows: The protein containing fractions were transferred to a dialysis bag and dialyzed overnight against 4 l buffer [20 mM Tris-HCl (pH 8.0), 100

mM NaCl, 50 mM KCl, 1 mM EDTA, 1 mM β-mercaptoethanol, 1μM PMSF, 0.1% triton-X-100, and 0.05% sarkosyl) containing 5 M urea. After overnight dialysis, the process was repeated with buffer containing 4M urea for 6 h, and every 6 h, the dialysis buffer was changed with a fresh buffer containing less urea until almost all of the urea was dialyzed out. The sample recovered was centrifuged at 10,000 rpm for 30 min at 4°C to remove the precipitates, if any. The sample was dialyzed against 2 l of the same buffer containing 50% glycerol but without any urea at 4°C overnight. The protein was stored at -80°C for further experimentation.

The fractions obtained while purifying the target protein from the soluble fraction were either dialyzed against a buffer containing 10 mM Tris-HCl (pH 7.4), 25 mM KCl, 1 mM DTT, 10 μM AEBSF [4-(2-Aminoethyl)-Benzenesulfonyl Fluoride Hydrochloride] and 50% glycerol or a buffer containing 50 mM Tris-HCl (pH7.4), 200 mM KCl, 10 mM EDTA, 1 mM DTT, and 50% glycerol, prior to storage.

HPLC Analysis

Some of the fractions obtained while purifying the target protein from the soluble fraction were sent to the Recombinant DNA/Protein Resource facility of Oklahoma State University for gel filtration analysis using HPLC. The proteins were eluted from the gel filtration column using either the elute buffer containing 200 mM imidazole or a buffer containing 50 mM Tris-HCl (pH 7.4), 200 mM KCl, 10 mM EDTA, 1 mM DTT, 0.1% triton-X-100 and 0.05% sarkosyl.

Extract Preparation for Methyltransferase assay

Total extracts from *E. coli* XL1- Blue MRA cells or those harboring various plasmids, *E.coli* cells that served as the expression hosts or those harboring plasmids, and *S. pneumoniae* cells with or without plasmids were obtained for use in the methyltransferase assay.

Total protein extract preparation from *S.pneumoniae*

Appropriate *S. pneumoniae* cells were grown in CATPG broth supplemented with 0.001% choline chloride to an OD_{550nm} of 0.2 (2x 10^9 CFU/ml) and 4x10^9 cells were centrifuged at 11,000 x g for 10 min at 4°C. The supernatant was discarded and the pellet was maintained at 0°C. The pellet was washed in 1 ml of buffer containing 10 mM Tris-HCl (pH 7.4), 25 mM KCl, and 1mM DTT and transferred to a microfuge tube at 0°C. This suspension was centrifuged at 14,000 rpm for 2 min, the pellet resuspended in 180 µl of the same buffer, and maintained at 0°C. To this, 20 µl of 10% triton-X-100 was added and the mixture was transferred to 37°C until clear. Immediately after lysis, the lysate was transferred to 0°C and was passed through a 21 gauge needle 15 times to shear the DNA. Then an equal volume of DNase free glycerol was added and the extract stored at -20°C until needed.

Total protein extract preparation from *E.coli* XL1-Blue MRA cells

Appropriate *E. coli* XL1-Blue MRA cells were grown overnight at 37°C with shaking. 50 ml of the cells were centrifuged at 11,000 x g for 10 min at 4°C. The pellet was resuspended in 500 µl of the buffer containing 10 mM Tris-HCl (pH 7.4), 25 mM KCl, 1mM DTT and 100µM PMSF, and maintained at 0°C. The suspension was transferred to a microfuge tube maintained at 0°C and then centrifuged at 11, 750 x g for 2 min. The pellet was resuspended in 300 µl of the same buffer and sonicated at 0°C ten times at 12 sec pulses. Then it was centrifuged at 11, 750 x g for 5 min and the supernatant was transferred to a fresh microfuge tube at 0°C and centrifuged once more. To the clear supernatant, an equal volume of DNase free glycerol was added and stored at -20°C until use.

Preparation of Soluble protein fraction from *E.coli* expression host

As mentioned in the large scale protein purification procedure, a glycerol stock of the appropriate expression host strain was streaked on agar plate with appropriate selection and incubated overnight at 37°C. An isolated colony was inoculated in 5 ml rich broth with appropriate selection and grown to an OD_{550nm} of 0.6 and then stored at 4°C overnight. One and eight-tenth milliliters of this culture was centrifuged at 11,750 x g for 1 min and resuspended in 500 µl of fresh rich broth. This resuspended mix was used to inoculate 90 ml of rich broth with appropriate selection and 2.5 mM betaine-HCl. The culture was incubated with shaking at 37°C until an OD_{550nm} of 0.6 was

reached. At this time point, 1mM IPTG was added to the culture and the cells were incubated at room temperature for 3 h with shaking. The cells were centrifuged at 8,000 rpm for 10 min and the pellet was resuspended in 5 ml of 1X binding buffer. To the resuspended cells, 0.1% triton-X-100 was added and stored at -20°C overnight. The cells were thawed in an ice-water bath and 0.05% Sarkosyl and 100 μM PMSF were added. The cells were sonicated at 0°C for 4 min with 20 sec pulses. The cells were centrifuged at 15,000 x g for 30 min and the supernatant was collected and filtered through a 0.45μ filter. To this filtered fraction an equal volume of DNase free glycerol was added and stored at -20°C until needed for the assay.

Protein Concentration Determination

Bradford Assay

The Bio-Rad Protein Assay, based on the method of Bradford (19), was the method of choice for determining protein concentration. A standard curve for the assay was plotted based on known concentrations of bovine serum albumin. Protein assay samples were prepared by mixing 10 μl of the protein fraction, 790 μl of ddH_2O, and 200 μl of the dye reagent concentrate. After vortexing thoroughly, the mixture was allowed to incubate at room temperature for at least 5 min. The samples were then transferred to disposable polystyrene cuvettes (Bio-Rad) and the absorbence measured using a spectrophotometer (Spectronic 1001, Milton Roy Company) set at 595 nm.

Comparison to the standard curve provided a relative measurement of protein concentration.

Methylation Activity Assay

The standard methylase assay mixture contained 5 μg of substrate DNA and 6 μg of protein extract in a total volume of 20-μl reaction buffer containing 50 mM Tris-HCl (pH 7.4), 10 mM EDTA, 0.5 mM DTT, and 1 μM [methyl-^{3}H]S-Adenosyl-L-methionine. After incubation at 37°C for 2 h, the assay mixtures were diluted to 250 μl with 10 mM Tris-HCl (pH 8.0), 1 mM EDTA, and extracted once with phenol: chloroform-isoamyl alcohol (25:24:1) and precipitated with ethanol. The precipitate was resuspended in 20 μl 10 mM Tris-HCl (pH 8.0), 1 mM EDTA. Determination of the incorporated radioactivity in the samples was done as described (110, 164).

CHAPTER IV

RESULTS

Tn*5253* was originally identified as a heterologous insertion in the chromosome of the clinical isolate of *S. pneumoniae* BM6001 (7, 177). Prior work in our lab has demonstrated this transposon to be a composite site-specific conjugative transposon consisting of two tranposons, Tn*5251* and Tn*5252* (7). Tn*5251* is 18 kb in size and belongs to the Tn*916* class of conjugative transposons (7). Tn*5252* is a 47.5 kb conjugative transposon that carries the chloramphenicol resistance gene. Since Tn*5252* is the prototype of complex conjugative transposons, our lab has focused on understanding the biology of the streptococcal conjugative transposon Tn*5252*. Previous studies have shown that Tn*5252* is a site-specific transposon. The nucleotide sequences of the right, the left, and the target site have been obtained (206). Comparison of these sequences has shown clearly that a 35 bp core sequence was adequate for the site-specific integration of the element (6). Studies to identify the transfer related regions of the transposon indicated that the ends of Tn*5252* were important for transfer (97). Preliminary work on the left end of the element showed the presence of a regulator (188), a DNA relaxase (189), an integrase and an excisionase (97) and more recently a gene conferring resistance to UV radiation (142).

Work on an internal region (located near the right end) has shown the presence of an operon like structure of many membrane-associated genes that have been implicated in transfer in other systems (1). The present work was undertaken to gain a better understanding of the right end of the element including the structural organization of the gene(s) present and their role in the transfer process. Fig. 2 presents the restriction map of Tn*5252* for 11 different restriction endonucleases, with the region of interest highlighted.

Cloning of the 3.27 kb *Eco*RI fragment in *E. coli*:

In order to construct the restriction map of Tn*5252*, Vijayakumar, *et al.* (208) introduced an *E. coli* vector, pVA891, at various sites in the transposon and generated various *E. coli* plasmids carrying flanking DNAs from the insertion mutants. One such plasmid, pVJ164 (207), contained a 3.27 kb *Eco*RI fragment forming the right junction fragment. The 3.27 kb *Eco*RI fragment was electroeluted from this plasmid for cloning purposes. Our initial attempts to clone this fragment in its entirety in several commonly used high copy number vectors such as, pBluescript SK^+ and pUC8 were unsuccessful suggesting that the passenger DNA might be unstable in high copy numbers. To overcome this problem, attempts were made to clone the fragment into pACYC184 and pVA891, which have a copy number of 18 (31) and 15 (126) respectively and these were also unsuccessful. This suggested that there might be some product encoded by

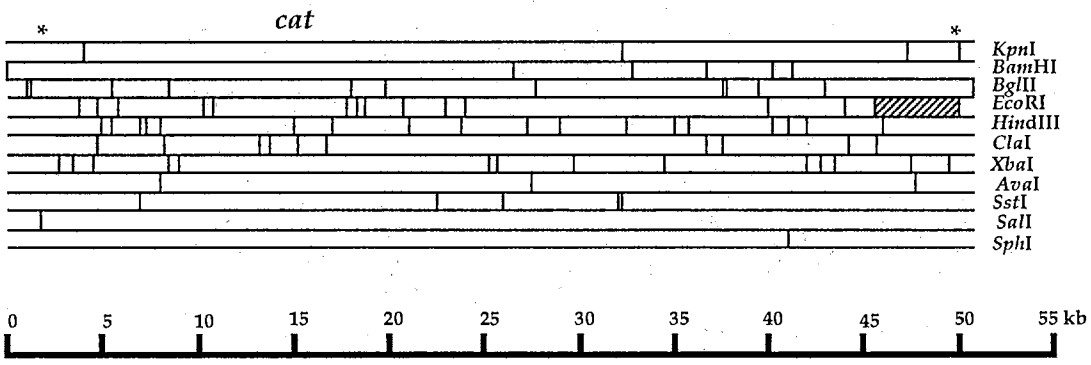


Figure 2. The restriction map of Tn*5252*, with the region of interest highlighted. The asterisks mark the end of the transposon. The location of the *cat* (chloramphenicol acetyl transferase) gene is indicated.

the passenger DNA that either might be toxic or might be unstable even in low copy numbers. The cloning attempts were repeated using a variety of *E. coli* hosts like HB101, JM109, C600, XL-1 and XL-1 Blue MRA (89). These attempts were also unsuccessful. In spite of our inability to clone this 3.27 kb *Eco*RI fragment in its entirety in *E. coli*, we were able to clone smaller fragments of this DNA in high copy number vectors. This strongly suggested that the passenger DNA encodes some gene product(s) that might be toxic to *E. coli* cells.

Cloning in *S. pneumoniae*

The inability to clone the 3.27 kb *Eco*RI fragment in *E. coli* led us to clone this fragment in *S. pneumoniae*. This cloning would be helpful in understanding the structural and functional characteristics of this region, as the right end of the transposon has been shown to be important for transfer (97). The vector chosen for this purpose was pLS1, a streptococcal 4.4 kb plasmid that has been shown to replicate in rec^+ *E. coli*. It has a copy number of 4 and 24 in *E. coli* and *S. pneumoniae*, respectively. The pneumococcal host strains used for cloning were Rx1 and SP1000 (Str^r Rx1::Tn*5252*). A recombinant plasmid, pSJ107 containing the 3.27 kb *Eco*RI right juncton fragment as insert was established in SP1000 but not in Rx1. This observation was in agreement with the report of Lopez, *et al.* (125) showing that the plasmid establishment in *S. pneumoniae* was increased 10 fold when the plasmid carried a segment of DNA homologous to the host chromosome.

Spontaneous Deletion of the cloned 3.27 kb *Eco*RI fragment in *E. coli*

To test our earlier hypothesis that the insert DNA specified a protein that might be toxic to *E. coli* cells, pSJ107 was introduced by transformation into competent *E. coli* cells and pLS1 was used as the control. As pLS1 can replicate only in rec^+ *E. coli* hosts, *E. coli* C600 and XL1-Blue MRA were chosen. The plasmid DNAs from single colony isolates were isolated. The results (Fig. 3) clearly indicated that pSJ107 suffered no deletion in XL1-Blue MRA as the 3.27 kb *Eco*RI insert was detected whereas in C600 cells several deleted forms of the plasmid were evident (lanes 5 and 6). The major difference in the genotype between the two host cells is that C600 is $McrA^-$ and XL1-Blue MRA is $McrA^-$, $McrBC^-$, Mrr^- and $hsdSMR^-$. It has been shown that restriction of methylated or methylase-encoding DNA by *E.coli* depends on the products of *mcr*A gene, the *mcr*BC operon and the *mrr* gene (143). The specificities of the McrA, McrBC and Mrr restriction activities for 5^mC containing DNA have been shown to be partially overlapping (143). Thus, depending on the modification pattern, the invading DNA may be subjected to restriction by one or several of these systems. Thus, restriction of pSJ107 by *E. coli* C600 indicated that the insert DNA contained a modification recognized by the host cells for restriction.

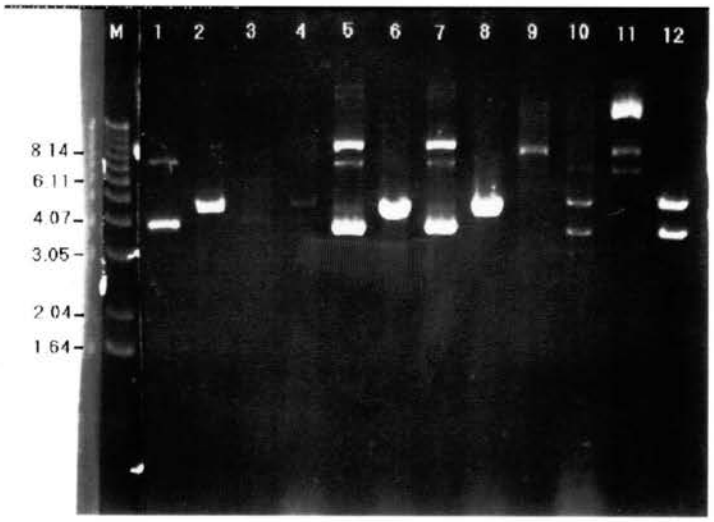


Figure 3. Restriction digest of the plasmids isolated from *E. coli* XL-1 Blue MRA and *E. coli* C600. Lane M, the 1 kb ladder. Nominal molecular masses are indicated. Lanes 1 and 2, pLS1 from *S. pneumoniae* (control); lanes 3 and 4, pLS1 from C600; lanes 5 and 6, pSJ107 from C600; lanes 7 and 8, pLS1 from XL1 Blue MRA; lanes 9 and 10, pSJ107 isolated from XL1 Blue MRA; lanes 11 and 12, pSJ107 from *S. pneumoniae* (control). Odd numbered lanes contain uncut plasmid DNA and even numbered lanes show *Eco*RI digested plasmids. pLS1 (the vector) is 4.4 kb in size and the insert is 3.27 kb in size.

Nucleotide sequence determination

To gain a better understanding of the possible protein products encoded by this segment of DNA, subclones containing segments of this fragment were generated. The restriction map of the 3.27 kb *Eco*RI fragment is shown in Fig. 4. The *Cla*I site is within the transposon and the right *Eco*RI is in the host chromosomal DNA about 400 bp away from the *att*R of Tn*5252* (206). The recombinant plasmids generated contained the 0.5kb *Eco*RI/*Xba*I fragment (pVJ430/431), the 1.8 kb *Xba*I fragment (pVJ419/442), the 1.1 kb *Kpn*I/ *Eco*RI fragment (pSJ114) and the 0.3 kb *Eco*RI/*Cla*I fragment (pSJ132) from this region. A nested set of deletion derivatives using Exonuclease III and S1 nuclease were generated from pVJ419, pVJ442 and pSJ114. The nucleotide sequence was obtained from all the subclones and the deletion derivatives. The 2.3-kb sequence from the right most *Eco*RI site to the left most *Xba*I site was obtained by Dr.Vijayakumar (personal communication). From these sequence-specific primers were designed for further sequencing reactions. The physical organization of the 3.5 kb fragment and the sequencing strategy used are shown in Fig. 5. The sequence between the *Xba*I and *Kpn*I site was obtained using pSJ107 as the template. The data obtained from all the subclones were manually assembled to form a contiguous sequence. Except for a 50-bp region to the left of the *Kpn*I site, the sequence was determined from both strands.

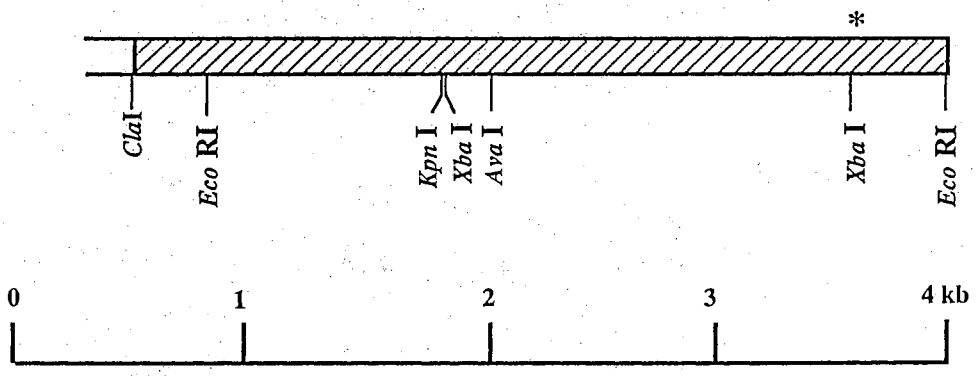


Figure 4. Restriction endonuclease map of the 3.5 kb fragment spanning the right end of Tn*5252*. The asterisk adjacent to the *Xba*I site marks the end of the transposon.

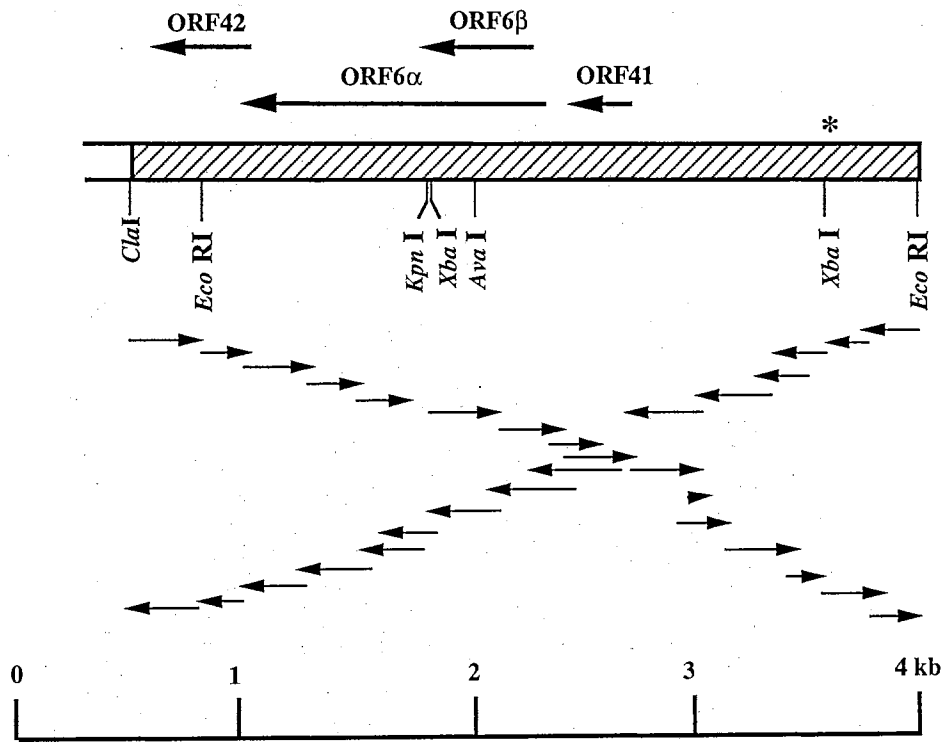


Figure 5. The restriction map of the 3.5-kb EcoRI fragment. The sequencing strategy and the different open reading frames (ORFs) are shown. For maximal confidence, both the strands of the DNA were sequenced. The arrows at the bottom show the sequencing strategy used. ORF6β is transcribed in an alternate reading frame and is within the ORF6α. The 5' end of ORF42 overlaps with the 3' end of ORF6α.

Structural analyses of the nucleotide sequence obtained

The DNA segment was 3565 bp long with a GC content of 36%. Computer analysis of the primary sequence data revealed the presence of four open reading frames (ORFs), designated ORF6α, ORF6β, ORF41, and ORF42. The predicted transcriptional map is shown in Fig. 5. The annotated nucleotide sequence of the DNA segment is shown in Fig. 6. All except ORF41 were preceded by an appropriately spaced consensus Shine-Dalgarno sequence. All were transcribed in the same orientation. ORF6α (nt 1704-3062) could specify a polypeptide of 453 amino acids of 50.7 kDa. ORF6β (nt1759-2277) localized within the region of ORF6α but in an alternate reading frame could encode a polypeptide of 173 amino acids with an apparent molecular mass of 20.1 kDa. ORF41 (nt 1320-1604) and ORF42 (nt 3046-3495) are presumed to encode polypeptides of 95 (11.3 kDa) and 150 amino acids (17.5 kDa) long, respectively. The 5'-end of ORF42 was also found to overlap with the 3'-end of ORF6α. Estimation of the net charge and isoelectric point (pI) showed the gene products of ORF6α and ORF6β to be basic polypeptides of 9.7 and 8.2, respectively. On the other hand, the estimated pIs of ORF41 and ORF42 polypeptides are 5.2 and 4.2 respectively. Upstream of ORF41, a region carrying 9-bp imperfect inverted repeats, IR2 (nt 1219- 1227 and nt 1246-1256), that could form a stem-loop structure with a ΔG = - 14.9 kcal/mol follwed by three nearly perfect 12-bp repeats, DR3 (nt 1269- 1281; nt 1284- 1295 and nt 1298- 1309), was observed. A promoter-like region showing consensus –10 and –35 sequences was present between the direct repeats

Figure 6. The nucleotide sequence of the 3.5 kb right junction fragment. The deduced amino acid sequences of the 4 ORFs (ORF41, ORF6α, ORF6β and ORF42) are also shown. The 10 base imperfect Inverted repeat (IR2) is present at the right most end of the transposon. ORF41 is flanked at the 5'end by a perfect 12 base repeat (IR1) and three nearly perfect 12 base pair Direct repeats (DR). The -10 and the -35 regions, the ribosome binding sites, and few of the restriction sites within the sequenced region are shown.

*Eco*RI

```
  1 GAATTCCTACTGATAATGAAGTAAGTTATGCTCTTATTTATCAGCAGGAAACTCTTCCTG

 61 CAACAGGTTCATCAACTTCTGAGCTTACAGCTTTAGGCCTATTAGCTGTTGGTAGTTTAG

121 TTCTTTTGGTTCATAATATGACGGGAACAGTTTTTTGCTCCCTCTGAAAAGTCATCATTT
                              DR1       DR1
                           ---------> --------->
181 GATGGCTTTTTCTATATAGGGTTAAAGATAGGCTAAAGATAGGGTAAAGGCTATCATCGA

241 CAAAATAAAGAAGGCATGATATAATATAAAGTAGATTTCTATGTCATAAAAACAAGAACT

301 ATTTGGACATCATTCATTTGAAAACTCTCTATGTTCAAACTATAGTAAAATAAAATAGGG

361 GATCTAAATCCTTGCTACGAGAAGGAAAAAAACTCAATGGCTACTATTCAATGGTTGGTC
                                                          IR1
    XbaI                                                 ---------->
421 TAGAAGATTTAAAATAATAAAACATTGTGAAGTAGAATAAATACTTAAAACCCTTGAAAT

        IR1
    <----------
481 TCAAGGGTTTTTATTGGTTTAGATTTTAATTATTCACATTGTTATATGTGCTTGATTGAG
                                  DR2       DR2
                               ------->  ------->
541 CAATAATCATCAATAAAAACATTTTTTTGGTGTTGAGTTTGGTGTTCAAATATTTATTAT

601 GTTACGGAAAGGAATTTCACAATCTTTCCGTTTTTTTATTTTCTGGAGGAAAAATGACAA

661 AAGAATTACAATCATCACGCTATATTGTCATTTCATTTTTAGTACGTGAAATGGGAATTG

721 ATATTGTTGAGGCCATATCTCTTATGGCTGAATTAGAAAAAAGTGGCTTGGTTCGCTTCG

781 AATCAAGTGGAGATTTAATACTCAAAGAACTTGGAGGAACGCTATGAAACGAATTACCGC

841 AAATCAATACCAAACTTACGAACGGTATTACAAATTACCTAAATTCTTTTTTGAGGGATG

901 AGAAATATATGGATATGAAACTAGAAGGTAAAGGTGGCTTATTCTATTTTAAAAGATCGT

961 TTAGAATTATCTCTCAGTCGTGGTTGGATAGATGAAGAGGGACGGTCTATTTAGTATTTC

1021 TAATTCTAAACTGATGAAGCTATTAGGTTGTTCGAAGTCAAAATTACTGTCCATCAAAAA

1081 AAACTCTTAAAGAATATGACTTAATTGATGAAGTTCAACAGTCTTCAAGTGAGAAAGGAA

1141 GACTAGCTAATAAGATTTATTTAGGGGAATTATCTTCTACCCCAGTAGCTAATTCAAACA
                           IR2                          IR2
                        --------->                   <---------
1201 GGCCATAGTGTTAAAAAAAGACTAGGGCAGGGTGAAAATGAAACGGCCCCCGTCTCACAT

                    -35                  -10
                    DR3          DR3          DR3
               ------------>  ----------->  ----------->
1261 TCAGCCCCTAGTGAGACTGAAGTTAGTGAGACTAATATAGTGAGACTGATTCTTTATTT
```

```
1320 ATT GAG GAT GAG GAG GAG AGG GAT ACT CAA CCT ATC TTG AAA AGA
ORF41> I   E   D   E   E   E   R   D   T   Q   P   I   L   K   R

1365 AAA GTA GAA AAG GTC ACA AAA TAT GAT CGA GAT TAT ATT TGG GGA
      K   V   E   K   V   T   K   Y   D   R   D   Y   I   W   G

1410 TTG GTG CAA GAT CAA TTT AGA CGA GAG GGG TTT TCT GAA ACA GCC
      L   V   Q   D   Q   F   R   R   E   G   F   S   E   T   A

1455 AGT GAA ATT GCT ATG ACT GAT TTT GAG AGA ATC TAT CAG TAT GCT
      S   E   I   A   M   T   D   F   E   R   I   Y   Q   Y   A

1500 CTT GAC AAT GTT CGC TTT GTT AGA CGA GCG GAA GTA CTT GCT GAA
      L   D   N   V   R   F   V   R   R   A   E   V   L   A   E

1545 TTT GTG TTT AAT GGC TTG TAT TCC GTT TGG AAC AAC CGT GTT AGA
      F   V   F   N   G   L   Y   S   V   W   N   N   R   V   R

1590 AAA GGA GGT GGT TAA ATGTCGCCAATCGAGTGGATATTAGTTATGCTTCATTGTC
      K   G   G   G   *

                                                    SD

1645 ATTTCAAGTGGAGTGACAGTTCTGACAATGATAGTCGATAAGAAAGGGGTGATAAAGAA

1704 ATG AGA TTT ATT GAT TTA TTT TCA GGT ATC GGT GGC TTT CGA CTA
ORF6α>M   R   F   I   D   L   F   S   G   I   G   G   F   R   L

             SD

1749 GGA ATG GAA AGT GTC GGA CAC GAG TGT ATT GGA TTT TGT GAG ATT
      G   M   E   S   V   G   H   E   C   I   G   F   C   E   I
          ORF6β>    V   S   D   T   S   V   L   D   F   V   R   L

1794 GAT AAA TTT GCT AGA GAA TCT TAT AAG TCC ATT TTT CAA ACG GAA
      D   K   F   A   R   E   S   Y   K   S   I   F   Q   T   E
       I   N   L   L   E   N   L   I   S   P   F   F   K   R   K

1839 GGA GAA ATT GAA TTT CAT GAC ATA CGA GAT GTT TCA GAT GAC GAA
      G   E   I   E   F   H   D   I   R   D   V   S   D   D   E
       E   K   L   N   F   M   T   Y   E   M   F   Q   M   T   N

1884 TTT AAA AAA CTT AGA GGG AAA GTC GAT GTC ATC TGT GGA GGA TTC
      F   K   K   L   R   G   K   V   D   V   I   C   G   G   F
       L   K   N   L   E   G   K   S   M   S   S   V   E   D   S

1929 CCT TGT CAA GCA TTT TCA ATC GCA GGA AGA CGA TTG GGA TTT GAA
      P   C   Q   A   F   S   I   A   G   R   R   L   G   F   E
       L   V   K   H   F   Q   S   Q   E   D   D   W   D   L   K

1974 GAT ACT AGA GGA ACT TTA TTC TTT GAA ATT GCT CGA GCG GCC AAA
      D   T   R   G   T   L   F   F   E   I   A   R   A   A   K
       I   L   E   E   L   Y   S   L   K   L   L   E   R   P   N
```

```
2019 CAA ATC CAA CCA CGT TTT CTT TTT CTT GAA AAT GTT AAA GGC CTA
      Q   I   Q   P   R   F   L   F   L   E   N   V   K   G   L
       K   S   N   H   V   F   F   F   L   K   M   L   K   A   Y

2064 CTC AAT CAC GAT AAG GGA CGG ACG TTC ACC ACA ATC CTT ACC ACG
      L   N   H   D   K   G   R   T   F   T   T   I   L   T   T
       S   I   T   I   R   D   G   R   S   P   Q   S   L   P   R

2109 CTT GAT GAA TTG GGG TTT GAT GTT GAG TGG CAG ATG CTT AAC AGT
      L   D   E   L   G   F   D   V   E   W   Q   M   L   N   S
       L   M   N   W   G   L   M   L   S   G   R   C   L   T   V

2154 AAG GAT TTT GGC GTT CCC CAA AAC AGA GAG AGG GTG TTT ATT ATC
      K   D   F   G   V   P   Q   N   R   E   R   V   F   I   I
       R   I   L   A   F   P   K   T   E   R   G   C   L   L   S

                XbaI                KpnI
2199 GGA CAT TCT AGA AAG AGA GGT ACC AGA CTC GGA TTT CCT TTC AGA
      G   H   S   R   K   R   G   T   R   L   G   F   P   F   R
       D   I   L   E   R   E   V   P   D   S   D   F   L   S   D

2244 CGA GAA GGT CAA GCA ACT AAC CCT GAG ACT TTA AAA ATA TTA GGG
      R   E   G   Q   A   T   N   P   E   T   L   K   I   L   G
       E   K   V   K   Q   L   T   L   R   L   *

2289 AAT TTG AAT CCA TCA AAA AGT GGA ATG AGT GGT AAA GTC TAT TAT
      N   L   N   P   S   K   S   G   M   S   G   K   V   Y   Y

2334 TCA GAA GGT CTT GCG CCA ACC TTA GTT CGT GGA AAA GGA GAA GGA
      S   E   G   L   A   P   T   L   V   R   G   K   G   E   G

2379 TTT AAA ATT GCG ATT CCT TGT ATG ACA CCA GAC AGA TTA GAC AAG
      F   K   I   A   I   P   C   M   T   P   D   R   L   D   K

2424 AGA CAA AAT GGT AGA CGT TTC AAG GAT AAT CAA GAG CCA ATG TTT
      R   Q   N   G   R   R   F   K   D   N   Q   E   P   M   F

2469 ACT TTA AAT ACT CAG GAT CGC CAT GGT ATT GTC GTT GTT GGA GAT
      T   L   N   T   Q   D   R   H   G   I   V   V   V   G   D

2514 TTA CCA ACT AGC TTT AAG GAA ACT GGT AGA GTC TAT GGA AGT GAG
      L   P   T   S   F   K   E   T   G   R   V   Y   G   S   E

2559 GGC TTA TCT CCA ACA CTG ACT ACG ATG CAA GGA GGA GAT AAA ATC
      G   L   S   P   T   L   T   T   M   Q   G   G   D   K   I

2604 CCC AAA ATA CTG ATT CCA GAA CCA ATC CAA TTT CTA AAA GTC CGA
      P   K   I   L   I   P   E   P   I   Q   F   L   K   V   R

2649 GAG GCA ACG AAA AAA GGA TAT GCT CAA GCA GAG ATT GGG GAT TCA
      E   A   T   K   K   G   Y   A   Q   A   E   I   G   D   S

2694 ATC AAT TTA GAA AGA CCA AGT TCT CAG CAT CGT CGT GGT AGA GTT
      I   N   L   E   R   P   S   S   Q   H   R   R   G   R   V
```

```
2739 GGA AAA GGT ATA GCT AAT ACG TTA ACA ACT AGT GGG CAA ATG GGA
       G   K   G   I   A   N   T   L   T   T   S   G   Q   M   G

2784 GTA GTA GTG GCT AGC TAT GAA GGA GAA GAT AAA CAA GTT TAC CAA
       V   V   V   A   S   Y   E   G   E   D   K   Q   V   Y   Q

2829 GTA GCT GGT GTT TTA ATA GAT GGA CAA TTT TAC CGT TTG AGA ATA
       V   A   G   V   L   I   D   G   Q   F   Y   R   L   R   I

2874 CGA CGA ATC ACT CCT AAA GAG TGT TTT CGA TTG CAG GGC TTT CCT
       R   R   I   T   P   K   E   C   F   R   L   Q   G   F   P

2919 GAT TGG GCT TTT GAA GCT GCT AGA AAA GTC TCT AGT AAT AGT CAG
       D   W   A   F   E   A   A   R   K   V   S   S   N   S   Q

2964 CTC TAT AAA CAA GCT GGT AAT AGT GTA ACC GTT CCT GTG ATT GCC
       L   Y   K   Q   A   G   N   S   V   T   V   P   V   I   A


                                                  SD

3009 GCA ATC GCA AAG AAA TTA AAA GAA GTA GAG GAA AAA GAT GAA AGC
       A   I   A   K   K   L   K   E   V   E   E   K   D   E
                                                 ORF42> M   K   A

3054 ATT AAA TAA A GAA TCA ATA CTA GAT TGT GAT GAA TTA GAA ACA
       I   K   *
         L   N   K     E   S   I   L   D   C   D   E   L   E   T

3097 GAA TTA CAT GAC GCA GAA ATC AAA CAG CTG GAT GAA CAA ATA TTT
       E   L   H   D   A   E   I   K   Q   L   D   E   Q   I   F

3142 TTG ATG CCC AAT TAT CCA TGT GAG TTT GAG GTG ACA TTT TTA GAT
       L   M   P   N   Y   P   C   E   F   E   V   T   F   L   D

3187 GAT TAC CAT AAA AAA CAC AAC TAC CCC CTA TTT TAC GAA TCC TAT
       D   Y   H   K   K   H   N   Y   P   L   F   Y   E   S   Y

                          EcoRI
3232 CTT CAA AAC ATT ATG GAA TTC CTT GAA AGT CAG GAT ATA AAG AAC
       L   Q   N   I   M   E   F   L   E   S   Q   D   I   K   N

3277 GGA GCT GAT GCC TTT GTA GAT GAT CAT CAG AAT CTT GTT TTT GTT
       G   A   D   A   F   V   D   D   H   Q   N   L   V   F   V

3322 TTA TAT GGA CAA GGC TAT AGA GCT GAG GGA AAA GAG GGA ATA CTT
       L   Y   G   Q   G   Y   R   A   E   G   K   E   G   I   L

3367 ACA ACC CAA GTA ACT GTA AAA GCT TTT GAT GAA GAC AAG AAA CCG
       T   T   Q   V   T   V   K   A   F   D   E   D   K   K   P

3412 ATT AAC TTC GCA AAT TTA TTA GAT TCC TTA ATC GTC TCA GAA TAT
       I   N   F   A   N   L   L   D   S   L   I   V   S   E   Y
```

```
3457 CAA ATG GAA CCG AAT CTT TGG GAG GTC TCC CAT GAT TGA TCTCTATC
      Q   M   E   P   N   L   W   E   V   S   H   D   *

                                                                ClaI
3504 TAAGTAAAAATAGCCAAAGAAATCAACTTCTTTTAGACTTCTTCCAAAACTATGGCATCGAT
```

(61, 160). Toward the very right end closer to the *att*R, another region of 12-bp inverted repats, IR1 (nt 467- 478 and nt 480- 491) with a $\Delta G = -16.6$ kcal/mol was found.

Sequence similarity search

The deduced amino acid sequences of the putative ORFs were analyzed for similarities to other protein sequences in the Genbank. Of the four ORFs, only ORF6α and ORF6β showed significant sequence similarities to other proteins. ORF6α showed similarities to a variety of prokaryotic 5^{m}C-cytosine methyltransferases. Multiple alignment of ORF6α with other 5^{m}C-cytosine methylases is shown in Fig. 7. The deduced amino acid sequence of ORF6α exhibited the same pattern of 10 conserved motifs that have been observed in all known 5^{m}C-DNA methylases (116, 143, 147, 202). The presence of the highly conserved sequence, PCXXXS within motif IV, containing the cysteine thought to form a transient, covalent complex with the cytosine to be methylated and the putative S-Adenosyl-L-Methionine (SAM, also called as AdoMet) binding site, (F,G)XGXG, in motif I provide compelling support for the suggestion that the ORF6α product could be a 5^{m}C-DNA methylase. The deduced amino acid sequence of ORF6α shares 72% similarity to a 5^{m}C-cytosine methylase present in the SPR phage of *Bacillus subtilis* (M.SPR). In addition to the 10 conserved motifs, significant sequence similarities between ORF6α and SPR phage methylase were also observed in the region located between the conserved regions VIII and IX. Fig. 8 shows the alignment of this variable region of ORF6α with that of the SPR phage methylase.

Figure 7. Multiple alignment of ORF6α with other 5^{m}C-cytosine methylases. The 10 conserved domains (I to X), found in all 5^{m}C-cytosine methylases are shown. The catalytic center is presumed to reside in the pro-cys dipeptide is highlighted. Amino acids represented in lower case indicate a gap. Gaps were introduced to maximize the similarity.

```
                1                                                  50
   ORF6α        MRFIDLFSGI GGFRLGMESV GHECIGFCEI DKFARESYKS IFQTEGEIEF
    M.SPR       LRVMSLFSGI GAFEAALRnv GYELVGFSEI DKYAVKSFCA IHNVDEQLNF
 M.ScrFIA       YKMIDLFAGI GGTRLGFHQT EkkSVFSSEI DKFAIKTYKA NFGDegDITK
  M.HpaII       FTFIDLFAGI GGFRIAMQNL GGKCIFSSEW DEQAQKTYEA NF...GDLPY
  M.SsoII       YRMIDLFAGI GGTRLGFHQT NAVNVVfsEW DKFAQKTYHA NYGddGDITK
   M.DsaV       LKFIDLFAGI GGMRIPFEEL GGKCVFSSEI DKHCQRTYEA NF...GEMPT
 M.phi3TI       LRVMSLFSGI GAFEAALRNI gyELIGFSEI DKYAIKSYCA IHNVSETLNV
  M.BsuFI       LTFIDLFAGI GGIRLGFEDK YTKCVFSSEW DKYAAQTYEA NY...GEKPH
  M.BsuRI       LRVMSLFSGI GAFEAALRNI gyELVGFSEI DKYAIKSYCA IHNADEQLNF
   M.HhaI       LRFIDLFAGL GGFRLALESC GAECVYSNEW DKYAQEVYEM NFGEKPEGDI
    M.Spβ       LRVMSLFSGI GAFEAALRNI gyELIGFSEI DKYAIKSYCA IHNVSETLNV
 M.HaeIII       MNLISLFSGA GGLDLGFQKA GFRIICANEY DKSIWKTYES NHS..AKLIK
    M.Dcm       FRFIDLFAGI GGIRRGFESI GGQCVFTSEW NKHAVRTYKA NHYCDPATHh
 M.EcoRII       FRFIDLFAGI GGIRKGFETI GGQCVFTSEW NKEAVRTYKA NWFNdrEVTL
 M.NgoPII       MKIISLFSGC GGLDLGFEKA GFEIPAANEY DKTIWATFKA NHPKTHLIE.
                          I                    II

                51                                                 100
   ORF6α        HDIRDVSDDE FKKLRGKVDV ICGGFPCQAF SIAGRRLGFE DTRGTLFFEI
    M.SPR       GDVSKID... .KKKLPEFDL LVGGSPCQSF SVAGHRKGFE DTRGTLFFQY
 M.ScrFIA       IDEKDIPDH. ........DI LVGGFPCQAF SQAGKKLGFD DTRGTLFFEI
  M.HpaII       GDITLEETKA F..IPEKFDI LCAGFPCQAF SIAGKRGGFE DTRGTLFFDV
  M.SsoII       IDEKDIPDHE .........I LVGGFPCVAF SQAGLKKGFN DTRGTLFFDI
   M.DsaV       GDITKLSADS IPY....HDL LLAGFPCQAF SQGGRKQGFQ DERGQLFFQV
 M.phi3TI       GDISKAKKDN IPYF....DL LTSGFPCPTF SVAGGRDGmk DERGTLFFET
  M.BsuFI       GDITKINEND IPDQ....DV LLAGFPCQPF SNIGKREGFA HERRNIIFDV
  M.BsuRI       GDVSKID... .KKKLPEFDL LVGGSPCQSF SVAGYRKGFE DTRGTLFFQY
   M.HhaI       TQVNEKTIPD H.......DI LCAGFPCQAF SISGKQKGFE DSRGTLFFDI
    M.Spβ       GDISKAKKDN IPYF....DL LTSGFPCPTF SVAGGRDGME YKCSNCSHEH
 M.HaeIII       GDISKISSDE FPK....CDG IIGGPPCQSW SEGGSLRGID DPRGKLFYEY
    M.Dcm       eDIRDItaEH IRQHIPEHDV LLAGFPCQPF SLAGVSKKNs dTQGTLFFDV
 M.EcoRII       SDKPEVPENd iNEHVPDHDV LLAGFPCQPF SLAGVSKKNs eAQGTLFFDV
 M.NgoPII       GDIRKIKEED FPE...EIDG IIGGPPCQSW SEAGALRGID DARGQLFFDY
                 III                          IV              V

                101                                                150
   ORF6α        ARAAKQIQPR FLFLENVKGL LNHDKGRTFT TILTTLDELG FDVEWQMLNS
    M.SPR       VETLKEKQPK FFVFENVKGL INHDKGNTLN VMAEAFSEVG YRIDLELLNS
 M.ScrFIA       ARIIKEKRPK AFLLENVKNL KTHDKGRTFK TILNTLEELD YEVHTALFKA
  M.HpaII       AEIIRRHQPK AFFLENVKGL KNHDKGRTLK TILNVLRegY FVPEPAIVNA
  M.SsoII       ARIIKEKKPH AFLLENVKNL LGHDKGRTFS IIKNTLEELN YTVYYNIFAA
   M.DsaV       AKILNDHRPQ AILLENVKGL RGHDKGRTLQ MILYVLEKLN YVVSWKIISA
 M.phi3TI       ALLAEEKKPK FVILENVKGL INSGNGQVLR IISETMNNIG YRIDLELLNS
  M.BsuFI       LRILKKKQPK MFLLENVKGL LTNDNGNTFR VILDNLKSLG YSVFYEVMDA
  M.BsuRI       IDTLKEKQPR YFVFENVKGL INHDKGNTLN IMAESFSEVG YRIDLELLNS
   M.HhaI       ARIVREKKPK VVFMENVKNF ASHDNGNTLE VVKNTMNELD YSFHAKVLNA
    M.Spβ       LITYEDYKKG VKCPKCEAVS KAKDERGTLF FETALLAEEK KPL...MVGH
 M.HaeIII       IRILKQKKPI FFLAENVKGM MAQRHNKAVQ EFIQEFDNAG YDVHIILLNA
    M.Dcm       VRIIDARRPA MFVLENVKNL KSHDQGKTFR IIMQTLDELG YDVAdpKIID
 M.EcoRII       ARIIRAKKPA IFVLENVKNL KSHDKGKTFK VIMDTLDELG YEVAdpKVID
 M.NgoPII       IRILKSKQPK FFLAENVSGM LANRHNGAVQ NLLKMFDGCG YDVTLTMANA
                          VI                              VII    VIII
```

```
                151                                                  200
   ORF6α        KDFGVPQNRE RVFIIGHSRK RGTRLGFPFR REGQATNPET LKILGNLNPS
    M.SPR       KFFNVPQNRE RLYIIGIrqKG KQRlsFNFR WTAQSAATKR LKDlgNINPS
M.ScrFIA        RDFGLPQNRE RIYIVGFDRK SISNYSDFQM PTPLQEKTRV GNILESVVDD
 M.HpaII        KNFGVPQNRE RIYIVGFHKS TGVN.SFSYP EPLDKIVTFA DIREEKTVPT
 M.SsoII        KDFGVPQNRE RIYIVGFNKE KVRNHEHFTF PTPLKTKTRV GDILEKSVDN
  M.DsaV        TDFNLPQKRE RIFIVGFQDK NNKNLIFDFP KPIELT.AKV GDLLEKEVDE
M.phi3TI        KFFNVPQNRE RVYIIGIRED LVENEQWVVG QKRNDVLSKG KKRLQEINIK
 M.BsuFI        QNFGLPQRRE RIVIVGFHPD LG..INDFSF PKGNPDNKVP INAILEHNPT
 M.BsuRI        KFFNVPQNRE RIYIIGVRED LIENDEWVVE KGRNDVLSKG KKRLKELNIk
  M.HhaI        LDYGIPQKRE RIYMICFRND LNIQ.NFQFP KPFELNTFVK DLLLPDSEVE
   M.Spβ        VDLKGHDAIK RVYSPEGLSP TLTTMGGGHR EPKIAEKQKE VRAV......
M.HaeIII        NDYGVAQDRK RVFYIGFRKE LNINYLPPIP HLIKPTFKDV IWDLKD....
   M.Dcm        GKHFLPQHRE RIVLVGFRRD LNLKADFTLR DISE...... ..........
M.EcoRII        GKHFLPQHRE RIVLVGFRRD LNIHQGFTLR DISRFYPEQ. ..........
M.NgoPII        KDYGVAQERK RVFYIGFRKD LEIKFSFPKG STVEDKDKIT LKDV......
                     VIII                                      TRD-I

                201                                                  250
   ORF6α        KSGMSGKVYY SEGLAPTLVR GKGEGFKIAI PCMTPDRLDK RQNGRRFKDN
    M.SPR       GNGMNGNVYN SSGLSPTITT NKGEGLKIAv pVLTPERGEK RQNGRRFKDD
M.ScrFIA        KYTISDKLWD GHQRRKTENK KNGKGFGYTL F......... ..........
 M.HpaII        KYYLSTQYID TLRKHKERHE SKGNGFGYEI I......... ......PDDG
 M.SsoII        KYTLSDALWN GHQRRKLVNA AAGKGFGYGL F......... ..........
  M.DsaV        KYTITDRMWE GHQNRKKAHR KRGNGFGFSL V......... ..........
M.phi3TI        SFNFKWPLQD TV........ .TKRLREILE DFVDEKYYLN EEKTKKLVEQ
 M.BsuFI        GYSISKRLQE SYLF...... .......... .......... ..........
 M.BsuRI        eEYVDEKYYL SEEKTSKLIE QIEKPKEKDV VFVGGINVGK rkQGNRVYDS
  M.HhaI        HLVIDRKDLV MTNQEIEQTT PKTVRLGIVG KG........ ..........
   M.Spβ        .......... .......... .......... ..LTPEREEK RQNGRRFKEN
M.HaeIII        .......... ..NPIPALDK NKTNGNKCIY PNHE...... ..........
   M.Dcm        .......... .......... .......... .......... ..........
M.EcoRII        .......... .......... .......... .......... ..........
M.NgoPII        .......... .......... .......... .......... ..........
                          TRD-I                          TRD-III

                251                                                  300
   ORF6α        QEPMFTLNTQ DRHGIVVVGD LPTSFKETGR VYGSEGLSPT LTTMQGGDKI
    M.SPR       GEPAFTVNTI DRHGVAV... .......... .......... ..........
M.ScrFIA        .......... .........N QDSEYTNTLS ARYYKDGSEI LIEQKNKNP.
 M.HpaII        IANAIVVGGM GRERNLVIDH RITDFTPTTN .......... ..........
 M.SsoII        .......... .........N ENSPYTNTIS ARYYKDGSEI LIEQKGSNP.
  M.DsaV        .......... .........N RNSSYTRTIS ARYYKDGSEV LVEQANKNP.
M.phi3TI        LGTAPLQKQE VREPLMVGHV DLKGHDAIKR VYSPEGLSPT LTTMGGGHRE
 M.BsuFI        .......... .......... .......... .......... ..........
 M.BsuRI        NGIATTLTSQ SVGGLGGQTS llKGHDAIKR VYSPDGVSPT LTTMGEGHRE
  M.HhaI        .......... .......... .....GQGER IYSTRGIAIT LSAYGGG...
   M.Spβ        GEPAFTVNTI DRHGVAI... .......... .......... ..........
M.HaeIII        .......... .....YFIGS YSTIFMSRNR VRQWNEPAFT VQASGRQCQ.
   M.Dcm        .......... .......... .......... ..CFPAQRVT LAQLLDPMVE
M.EcoRII        .......... .......... .......... ...RPSFGEL LEPVVDSKYI
M.NgoPII        .......... ..........  .......... .......... ..........
                     TRD-III
```

```
                301                                                  350
   ORF6α        PKILIPEPIQ FLKVREATKK GYAQAEIGDS INLERPSSQH RRGRVGKGIA
    M.SPR       .......... .......... .......... .......... ..........
M.ScrFIA        .......... .......... .......... .......... ..........
 M.HpaII        .......... ..........  ........... ......... ..........
 M.SsoII        .......... .......... .......... .......... ..........
  M.DsaV        .......... .......... .......... .......... ..........
M.phi3TI        PKI....... .......... AEKQKEVRAV LTPEREEKRQ NGRRFKENGE
 M.BsuFI        .......... ........KK DDGKPQIVDF RCTYQVNTLV ASYHKIQRLT
 M.BsuRI        PKIAVEYVGN INPSGKGMND QVYNSNGLSP TLTTNKGEGV KISVPNPEIR
  M.HhaI        .......... .......... .......... .......... ..........
   M.Spβ        .......... .......... .......... .......... ..........
M.HaeIII        ...LHPQAPV MLKVSKNLNK FVEGKE.... .......... ..........
   M.Dcm        AKYILTPVLW KYLYRYAKKH QARGNGFGYG MVYPNNPQSV TRTLSARYYK
M.EcoRII        LTPKLWEYLY NYAKKHAAKG NGFGFGLVNP ENKESIARTL SARYHKDGSE
M.NgoPII        ....IWDLQD TAVPSAPQNK TNPDAVNNNE YFTGSFSPIF MSRNRVKAWD

                351                                                  400
   ORF6α        NTLTTSGQMG VVVASYEGED KQVYQVAGVL IDGQFYRLRI RRITPKECFR
    M.SPR       .......... .......... .......... ..GEYPKYRI RRLTPLECFR
M.ScrFIA        .......... .......... .......... .......... RKITPREAAR
 M.HpaII        .......... .......... .......... IKGEVNREGI RKMTPREWAR
 M.SsoII        .......... .......... .......... .......... RKITPREASR
  M.DsaV        .......... .......... .......... .......... RVLTPRECAR
M.phi3TI        PAFTVNTIDR HGVAI..... .......... ..GEYPKYKI RKLSPLECWR
 M.BsuFI        GTFVKDGETG .......... .......... .........L RLFSELELKR
 M.BsuRI        PVLTPEREEK RQNGRRFKED DEPAFTVNtv AIGEYPKYRI RKLTPLECWR
  M.HhaI        .......... .......... ......IFAK TGGYLVNGKT RKLHPRECAR
   M.Spβ        .......... .......... .......... ..GEYPKYKI RKLSPLECWR
M.HaeIII        .......... .......... .......... .......HLY RRLTVRECAR
   M.Dcm        DGAEILIDRG WDMATGEKD. .........F DDPLNQQHRP RRLTPRECAR
M.EcoRII        ILIDRGWDMA TGETDFANEE NQAH...... .........RP RRLTPRECAR
M.NgoPII        EQGFTVQASG RQCQLHPQAP KMEKHGANDY RFAAGKETLY RRMTVREVAR
                                                              IX

                401                                                  452
   ORF6α        LQGFPDWAFE AARKVSSNSQ LYKQAGNSVT VPVIAAIAKK LKEVEEKDES IK
    M.SPR       LQAFDDEDFE KAFAaiSNSQ LYKQTGNSIT VTVLESIFKE LIHTYINKES E.
M.ScrFIA        LQGFPENFII PV....SDTQ AYKEFGNSVA VPTIHAIAEK MLEVLEKSKK ..
 M.HpaII        LQGFPDSYVI PV....SDAS AYKQFGNSVA VPAIQATGKK ILEKLGNLYD ..
 M.SsoII        LQGFPSDFII PV....SDTQ AYKQFGNSVA VPVINAIAEK IISTLDS... ..
  M.DsaV        LQGFPESFVI PV....SDCQ AWRQFGNSVP VSVIRAIAQK MLSYIDLTEQ QK
M.phi3TI        LQAFDDEDFE KAFAaiSNSQ LYKQAGNSIT VSVLESIFQE LIHTYVNKES E.
 M.BsuFI        LMGFPVDF.. ..KVPVSRTQ MYRQFGNSVA VPMIKAVAGA MKERLLLAEM ..
 M.BsuRI        LQAFDEEDFE KALSviSNSQ LYKQAGNSIT VTVLESIFKE LIHTYVNEES E.
  M.HhaI        VMGYPDSY.. ..KVHPSTSQ AYKQFGNSVV INVLQYIAYN IGSSLNFKPY ..
   M.Spβ        LQAFDDEDFE KA........ .......... .......... .......... ..
M.HaeIII        VQGFPD...D FIFHYESLND GYKMIGNAVP VNLAYEIAKT IKSALEICKG N.
   M.Dcm        LMGFEAPGEA KFRIPVSDTQ AYRQFGNSVV VPVFAAVAKL LEPKIKQAVA LR
M.EcoRII        LMGFEKVDGR PFRIPVSDTQ SYRQFGNSVV VPVFEAVAKL LEPYILKAVN AD
M.NgoPII        IQGFPD...N FKFIYQNVND AYKMIGNAVP VNLAYEIAAA IKKTLER... ..
                 IX                          X
```

Figure 8. Alignment of variable region of ORF6α with that of the methyltransferase encoded by *Bacillus subtilis* phage SPR. Vertical bars represent identical amino acids and dots show conserved substitutions. Gaps were introduced to maximize the similarity. The overall similarity between ORF6α and M.SPR is 72% and identity is 52%. Each of the target recognition domains (TRDs), TRD-I, TRD-II, TRD-III, of the phage SPR methylase have been shown by others (66) to be responsible for the recognition of target regions, *Eco*RII, *Msp*I, and *Hae*III restriction endonuclease sites correspondingly for methylation. Regions similar to TRD-I and TRD-III, are present in ORF6α.

The SPR phage methylase has been shown (66) to contain three target-recognizing domains (TRDs), that are responsible for recognizing three different target sequences. TRDI has been shown to recognize the *Eco*RII site, TRDII, the *Msp*I site and TRDIII, the *Hae*III site (9, 204, 214). Based on the alignment, ORF6α was found to have regions corresponding to TRDI and TRDIII, but not TRDII. Fig. 9 shows the alignment of the TRDI of ORF6α with the corresponding domains of the multi-specific methylases SPR, ϕH2, ϕ3TI, ϕSpβ, and ϕρ11sI. The TRDs of M.*Eco*RII, M.Dcm both of which recognize the sequence CCWGG, the TRDs of M.*Sso*II, M.*Nla*X and M.*Scr*FI all of which recognize the sequence CCNGG, and the TRD of M.*Dsa*V which recognizes the sequence CCTGG are also aligned. As evident from the figure, the multi-specific methylases of the *B. subtilis* phages and the ORF6α methylase group together and share a common domain structure whereas the mono-specific methylases form a separate group and share a common structure. Similar reports have been published (143,203). This structural difference between the multi-specific and the mono-specific methylases recognizing the same target site indicates that there might be distinct requirements for mono-specific and multi-specific methylases. Walter, *et al.* (212) have shown that the multi-specific methylases could not be endowed with novel specificities by integrating TRDs of mono-specific methylases. This suggests that in the case of multi- and mono-specific methylases, the interaction of the TRDs with their core sequence is different. ϕH2, and ϕρ11 have been shown not to be able to methylate the *Eco*RII site (CCWGG) in spite of sharing significant amino acid similarity with the

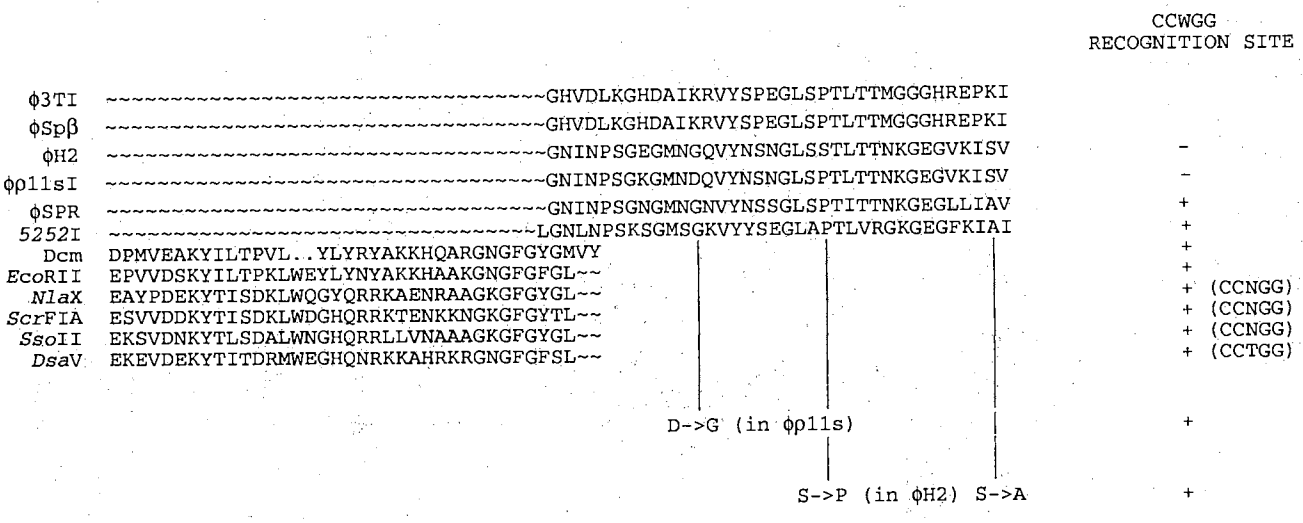


Figure 9. Alignment of TRDI of ORF6α (*5252*I) with the corresponding domains of other multi-specific and mono-specific methylases. The substitutions made in the pseudo domain of ϕρ11sI and ϕH2 that made them active are also indicated (112). + indicates the ability of an enzyme to recognize and methylate CCWGG and – indicates the inability. ? indicates uncertainty recognition and methylation. Domains responsible for recognizing other sites (CCNGG or CCTGG) are indicated.

TRDI (*Eco*RII target reconizing domain) of ϕSPR (112). These inactive domains (pseudo domains) have been activated by site-specific mutations. Based on the alignment and the activating mutations made in the case of ϕρ11 and ϕH2, the TRDI of ORF6α could be a functional *Eco*RII domain. Fig. 10 shows the alignment of TRDIII of ORF6α with the corresponding TRDs of multi-specific phage methylases, and mono-specific methylases that recognize the sequence GGCC. As in the case of TRDI, here also the mono-specific methylases from a separate group compared to the multi-specific methylases. The alignment, shown in Fig. 11, of a region next to TRDIII of ORF6α with that of other multi-specific phage methylases revealed a significant sequence the entire *Fnu*4HI domain but only a portion of it. This partial sequence similarity indicated that the sequence recognized by ORF6α might be similar to the sequence GCNGC, but not the same. In the case of M.ρ11sI, the *Fnu*4HI domain has been shown to be non-functional. The amino-acid sequence of the *Fnu*4HI recognizing domain of ϕH2I and ϕρ11sI are identical except for one amino acid difference (112) and conversion of this Glutamic acid to Glycine by site-specific mutation has been shown to restore the methylation potential of ϕρ11sI. In the case of ORF6α, the homologous position is occupied by Glycine, which could mean that the domain is functional.

		GGCC RECOGNITION
ϕ3T	AVLTPEREEK RQNGRRFKEN GEPAFTVNTI DRHGVAI	+
ϕSpβ	AVLTPEREEK RQNGRRFKEN GEPAFTVNTI DRHGVAI	+
ϕH2	PVLTPEREEK RQNGRRFKED DEPAFTVNTI DRHGVAI	+
ϕρ11sI	PVLTPEREEK RQNGRRFKED DEPAFTVNTI DRHGVAI	+
ϕSPR	PVLTPERGEK RQNGRRFKDD GEPAFTVNTI DRHGVAV	+
ORF6α	PCMTPDRLDK RQNGRRFKDN QEPMFTLNTQ DRHGIVV	?
*Bsp*RI	PYFTGSYSTI FMSRNRKKKW TDQSFTIQAS GRQAPIH	+
*Bsu*RI	PYFTGSYSST YMSRNRKKSW .DQSFTIQAS GRQAPLH	+
*Hae*III	EYFTGSYSTT FMSRNRVRQW NEPAFTVQAS GRQCQLH	+
*Ngo*PII	EYFTGSFSPI FMSRNRVKAW DEQGFTVQAS GRQCQLH	+

Figure 10. Alignment of the TRDIII of ORF6α with the corresponding domains of multi-specific and mono-specific methylases. + indicates the ability to recognize and methylate GGCC, – indicates the inability and ? indicates uncertainty.

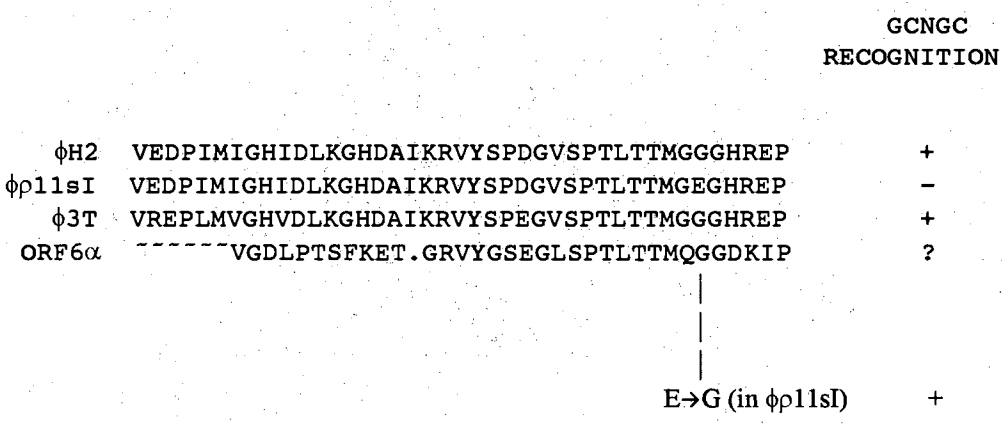


Figure 11. Alignment of a region next to TRDIII of ORF6α with the corresponding domains of multi-specific methylases. + indicates the ability to recognize and methylate the GCNGC site and – refers to the inability to do so. A site-specific mutation made in ϕρ11sI to make the domain active is also shown (112). ? refers to the uncertainty in the recognition site.

ORF6β homology

ORF6β showed significant similarity to β sub-unit of M.*Eco*HK311. Methylation of a substrate DNA by M.*Eco*HK311 has been shown to require two proteins, α and β. One of the proteins has been shown to be translated in an alternate reading frame within the other. Nine of the ten conserved motifs found in the 5^{m}C-methylases were found in the M.*Eco*HK311α polypeptide. Motif IX was found to be present in the β polypeptide (122). Fig. 12 shows the alignment of ORF6β with the β sub-unit of M.*Eco*HK311. However, the region of similarity between the two polypeptides does not span the conserved motif IX and no conserved motif was detected in ORF6β.

Protein Modeling

The crystal structure of four type II methyltransferases, M.*Hha*I (34), M.*Hae*III (155), M.*Taq*I (108) and M.*Pvu*II have been resolved (129). Of these four, M.*Hha*I and M.*Hae*III are both 5^{m}C-cytosine methyltransferases, M.*Taq*I a 6^{m}N-adenine methylase and M.*Pvu*II is a 4^{m}N-cytosine methylase. Presence of a bilobal structure has been established by comparing the structures of these 4 DNA methylases (129). One lobe consists of the catalytic domain with both active site for methyl transfer and the AdoMet binding site and the other lobe contains the target recognition domain. Comparison of small molecule methylases, catechol-O-methylase, DNA methylases, and other

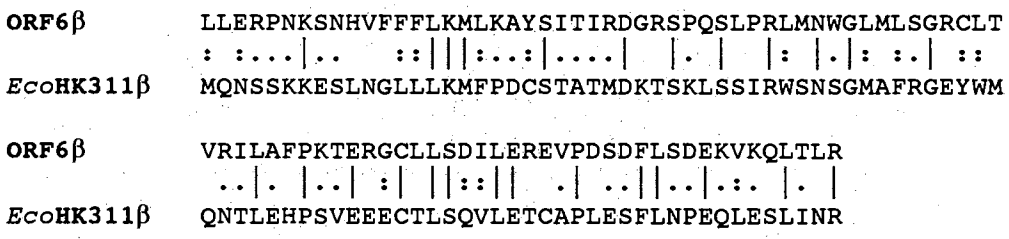


Fig. 12. Alignment of the predicted protein product of ORF6β with the β-subunit of M.*Eco*HK311. Vertical bars represent identity and dots represent similarity. The overall similarity is 46% and identity is 29%.

methylases has revealed a similar three-dimensional folding of the catalytic domain in all these enzymes (129). Significant similarity between the putative catalytic domain of ORF6α and M.*Hha*I and M.*Hae*III was observed (Fig. 7). Hence homology-based protein modeling was done using these known structures as templates. Due to the lack of homology between the target recognition domain of ORF6α and the other methylases, this domain has not been modeled. The model thus developed is shown in Fig. 13. The catalytic domain is dominated by α/β sheet structure that is very similar to those of other DNA methylases. The main feature is a seven stranded β-sheet (6↓ 7↑ 5↓ 4↓ 1↓ 2↓ 3↓) formed by five parallel β-strands and an antiparallel β-hairpin. The β-sheet is flanked by six parallel α-helices, three on each side. The AdoMet binding site is located at the C-terminal ends of β1 and β2 and the active site is at the C-terminal ends of strands β4 and β5 and the N-terminal end of β7.

In vitro Transcription and Translation

In order to analyze the protein products encoded by the 3.27-kb *Eco*RI fragment, pSJ107 and pLS1 were added *in vitro* to an *E. coli* S30 extract. The radioactively labeled polypeptides were resolved on a 12% polyacrylamide gel under denaturing conditions. As shown in Fig. 14, four polypeptides of apparent molecular masses of 46, 42, 21 and 19.5 kDa unique to pSJ107 were identified. The molecular masses of two of these, 46 and 21 kDa, were in reasonable agreement with the molecular masses of ORF6α and ORF6β proteins deduced from the nucleotide sequence data. The

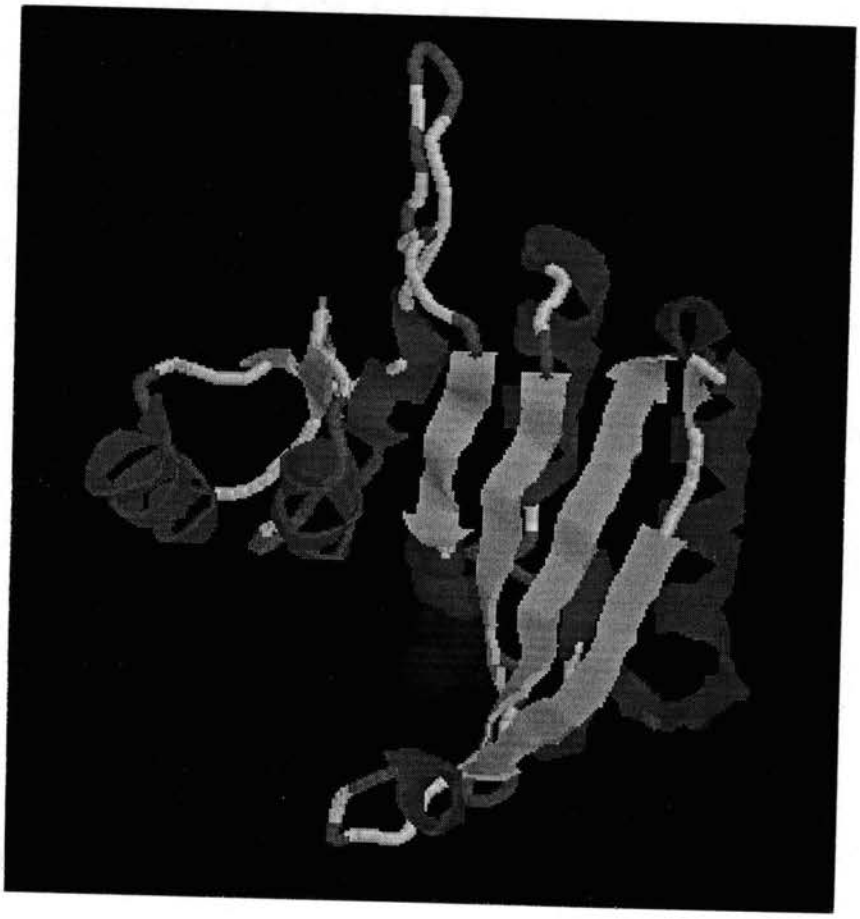


Figure 13. Homology-based modeling of the catalytic domain of ORF6α. The modeling was done using the SWISS-MODEL and is based on the structure of M.*Hha*I methylase. The Pro-Cys dipeptide that forms the active site is shown using the spacefill model. The β-sheets and α-helices are also shown in the model.

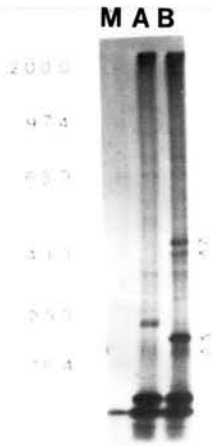


Figure 14. *In vitro* Transcription and Translation using the *E.coli* S30 Extract with pLS1 and pSJ107 as templates. pLS1 and pSJ107 plasmid DNA were added to *E. coli* S30 extract and the protein products made were electrophoresed on a 12% SDS-PAGE. Lane (A), pLS1; (B), pSJ107 and (M). Molecular weight standard, the sizes of the molecular weight markers indicated are in kilodaltons. The 46 kDa band and the 21 kDa band in lane (B) corresponds to ORF6α and ORF6β proteins respectively.

expected sizes for ORF6α and ORF6β were 50.7 kDa and 20.1 kDa respectively. It was assumed that the 42-kDa polypeptide was the result of fusion between the amino terminal portion of the truncated ORF42 and the ORF-D protein of the vector (109). It is not clear at present whether the 19.5 kDa band was the product of proteolysis or translation from an alternate start site. No product resembling the predicted polypeptide of ORF41 (expected size of 11.3 kDa) could be observed due to the presence of radioactivity at the expected position in the control sample containing the vector.

Construction of pSJ117

In order to mutate ORF6 in pSJ107 and to check its effect on the methylase activity, the central four bases within the unique *Kpn*I site in pSJ107 was removed by Klenow treatment. The plasmid, designated pSJ117, was generated in *E.coli* XL1-Blue MRA. This plasmid was expected to produce a truncated methylase protein due to the frame shift mutation introduced. Efforts to generate pSJ117 from *E.coli* C600 were unsuccessful, indicating that the truncated methylase produced from the plasmid was still active. To confirm the loss of the *Kpn*I site, pSJ117 was subjected to restriction digestion. Fig. 15 shows the restriction digestion of pSJ107 and pSJ117 along with pBluescript SK^+. The nucleotide sequence of pSJ117 was also obtained using a newly designed synthetic primer, which also confirmed the loss of the central four bases of the *Kpn*I site.

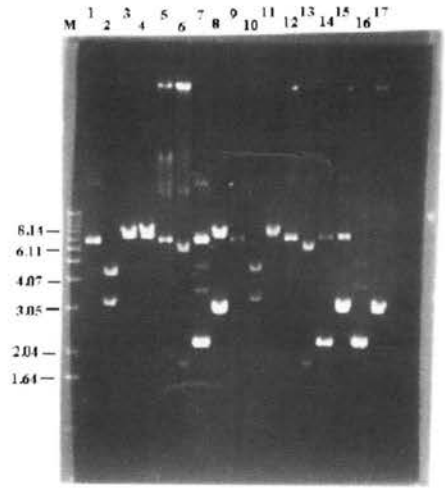


Figure 15. Restriction digest showing the loss of the *Kpn*I site in pSJ117.

Undigested and digested DNA were run on a 0.8% agarose gel. pSJ107 and pBluescript SK^+ digested with various enzymes served as the control. The molecular masses (lane M) in kilobases are indicated. Lanes 1, 5, pSJ107 uncut; lane 9, pSJ117 uncut. pSJ107 and pSJ117 were digested with *Eco*RI (lanes 2, 10 respectively); *Hind*III (lanes 3, 11 respectively); *Kpn*I (lanes 4,12 respectively); *Xba*I (6, 13 respectively). pBluescript SK^+ was mixed with pSJ107 (lane 7) and pSJ117 (lane 14) and the mix was digested with *Kpn*I (lane 8 and 15 respectively). pBluescript SK^+ uncut (lane 16) and *Kpn*I digest (lane 17).

In vitro Assay for Methyltransferase activity

To demonstrate directly the DNA methylase activity, crude extracts from *E. coli* and pneumococcal cells with or without pSJ107 were used in a reaction buffer containing [*methyl*-^{3}H] S-adenosyl methionine (AdoMet) and *E. coli* GM2163 chromosomal substrate DNA. Background numbers obtained in the absence of extracts were subtracted. The results are presented in Table V. Based on the incorporated radioactivity retained on the filters, it was evident that the passenger DNA in pSJ107 conferred DNA methylase activity to *E. coli* and pneumococcal cells harboring it. Methylase activity of extracts of *E. coli* cells carrying pSJ117 was found to be 30-fold higher. This result was not altogether surprising as the amino terminal half of 5^{m}C-cytosine methylases are known to methylate random sequences in the absence of target recognizing domains at the carboxyl end (67). Interestingly pneumococcal strain Rx1 was observed to carry abundant methylase activity that did not significantly increase when copies of Tn*5252* methylase genes were added. This could possibly be due to down-regulation of expression in the native host, especially in the light of the various repeats found upstream of the gene. As Fig. 16 shows, the incorporation of radioactivity was dependent on the concentration of the extract added.

Table V. Incorporation of [^{3}H] AdoMet to Nonmethylated *E. coli* GM2163 DNA

EXTRACT	cpm
***E. coli* XL1 Blue MRA**	19
***E. coli* XL1 Blue MRA (pLS1)**	5
***E. coli XL1* Blue MRA (pSJ107)**	521
***E. coli XL1* Blue MRA (pSJ117)**	15249
***S. pneumoniae* Rx1**	1871
***S. pneumoniae* Rx1 (pLS1)**	2661
***S. pneumoniae* Rx1 (pSJ107)**	2160

Six microgram of protein extract and 5μg of substrate DNA (GM2163 chromosomal DNA) were used in each reaction under the conditions described. The values shown are the average of two independent experiments.

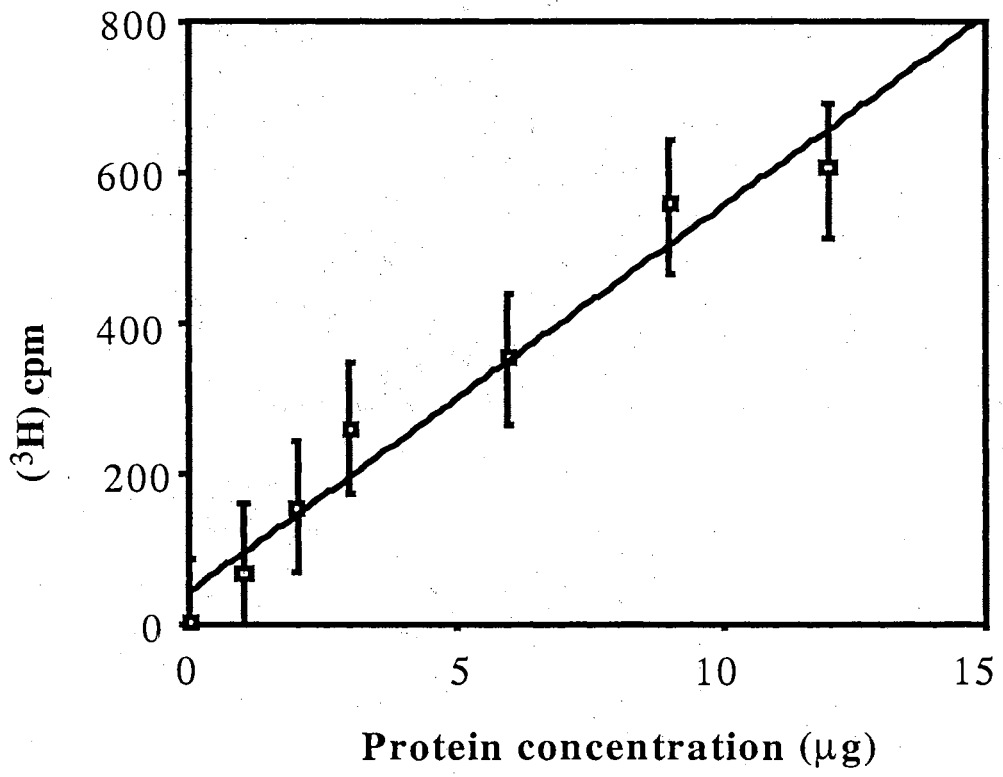


Figure 16. Graph showing the concentration dependence of the methylase activity. Extract from *E. coli* XL1- Blue MRA cells carrying pSJ107 was used for the assay. The background radioactivity incorporated in the absence of the methylase was subtracted.

Over expression and purification of the methylase

To purify the methylase protein and study its biochemical properties, ORF6α was cloned into the expression vector, pMal-2.

The pMal-2 system

The pMal-2 vectors ($pMalC_2$ and $pMalP_2$) express the *mal*E gene which encodes the maltose binding protein (MBP) of *E. coli* fused to the *lac*Zα gene. Seven unique restriction sites are available between *mal*E and lacZα. The gene of interest is cloned in the multiple cloning site and the insertion inactivates the *lac*Zα resulting in no α-complementation. The vectors have a strong *tac* promoter and the *mal*E translation initiation signals to give high-level expression of the cloned gene (4, 54) and one step purification of the fusion protein using MBP's affinity for maltose (96). The vectors carry the *lac*I^q gene coding for the Lac repressor which keeps the expression from P_{tac} low in the absence of IPTG. They also contain the recognition site for the site-specific protease, factor Xa, just 5' of the multiple cloning sites. Thus, after purification the MBP can be cleaved off the fusion protein. The over-expressed protein, in the case of $pMalC_2$, found in the cytoplasmic fraction constitutes 20-40% of the total cellular protein. Due to the presence of the signal peptide, in the case of $pMalP_2$, the fusion protein is transported to the periplasm and it constitutes 5-10% of the total cellular protein.

Cloning ORF6α in pMalC$_2$

The template used for amplifying ORF6α by PCR was pSJ107 linearized with *Hin*dIII. PCR was performed as described previously using the primers MET-F (5'-ATGAGATTTATTGATTTATTTTCAGGT-3') and MET-R (5'-TTTATTTAATGCTTTCATCTTTTTCC-3'). The amplified product (1.4-kb) was gel purified and ligated to *Xmn*I digested pMalC$_2$. *E. coli* JM109 cells were transformed with the ligation mix and white colored transformants were selected on LB agar with ampicillin and X-gal. These transformants were then replica-plated on LB agar containing ampicillin and X-gal or ampicillin, X-gal and IPTG. Colonies that remained white in the plates with X-gal and IPTG were chosen and the corresponding colonies in X-gal plates were used for making a mini plasmid preparation. The plasmids were digested with *Xba*I to check the orientation of the insert, as there were two sites for this enzyme, one in the insert and one in the vector. The nucleotide sequences of three independent recombinant plasmids were also obtained. The restriction digest and the nucleotide sequence information revealed the presence of the insert in the wrong orientation. The sequence also revealed the presence of 7-8 bp deletions. This was assumed to be due to the 3'-5' exonuclease activity of DeepVent polymerase. To avoid this problem, the dNTPs concentration was increased from 250 μM to 625 μM, 1.25 mM and 2 mM and three independent PCR reactions were set up. After amplification, a sample was electrophoresd on the gel and it was observed that the non-specific amplification was reduced with an increase in the dNTP concentration. It was also noted that the specific product amplification reduced with the increase in dNTPs. The PCR

products obtained from the reaction using 1.25mM and 2mM dNTPs were chosen for further work. The amplified fragment was gel purified and used in ligations with $pMalC_2$ digested with *Xmn*I. The recombinant plasmids obtained from three of the clones were restriction digested and also sequenced. Both restriction digestion and the nucleotide sequecing revealed that the insert was in the wrong orientation. However, unlike the previous experiment, no deletion in the insert DNA was observed. One such plasmid was designated pSJ133. The observation that the methylase gets cloned in the reverse orientation in $pMalC_2$ indicated that the leaky expression of the MBP-methylase fusion product might be toxic to the host even if the expression is at a very low-level, in the absence of induction by IPTG.

Cloning ORF6α in $pMalP_2$

The inability to clone the methylase in $pMalC_2$ led me to try and clone the methylase in $pMalP_2$ as it would be exported into the periplasm. The PCR product was gel purified and ligated to $pMalP_2$ digested with *Xmn*I. The ligation mix was used to transform JM109, DH5α and XL1-Blue MRF'. The recombinant plasmids obtained from JM109 and DH5α contained deletions but the plasmid obtained from MRF' had the methylase in the correct orientation. This was checked by restriction enzyme digestion and confirmed by nucleotide sequencing. This plasmid was designated pSJ136 and was used for protein purification pilot experiments.

Protein purification using pSJ136 (pilot)

Before the large-scale purification using pSJ136, a pilot experiment was done to check the level and the duration of induction. The cells were grown to an OD_{600nm} of 0.5 (2×10^8 CFU/ml) and were induced with IPTG for 3 h. Samples for SDS-PAGE were taken before the start of induction and after every hour during induction. After three h, crude extract and osmotic shock fluid were collected. The crude extract was also mixed with the amylose resin, to check the amount of the induced MBP-methylase fusion protein bound to the resin. The bound fusion protein was released with maltose and was run on SDS-PAGE. As evident from the result shown in Fig. 17, the induction of the fusion protein, the expected size of which is about 92 kDa (42 kDa of the MBP with 50 kDa of the methylase), was inadequate to purify the protein in large quantities. Therefore, the methylase was cloned into another expression vector.

The pET system

The gene of interest is cloned into one of the pET plasmids and placed under the control of a strong bacteriophage T7 transcription signal. The expression of the gene of interest is achieved by providing the host cells with a source of T7 RNA polymerase.

The T7 RNA polymerase gene is under the control of the *lac*UV5 promoter in the chromosome of the host and can be induced by the addition of IPTG. After

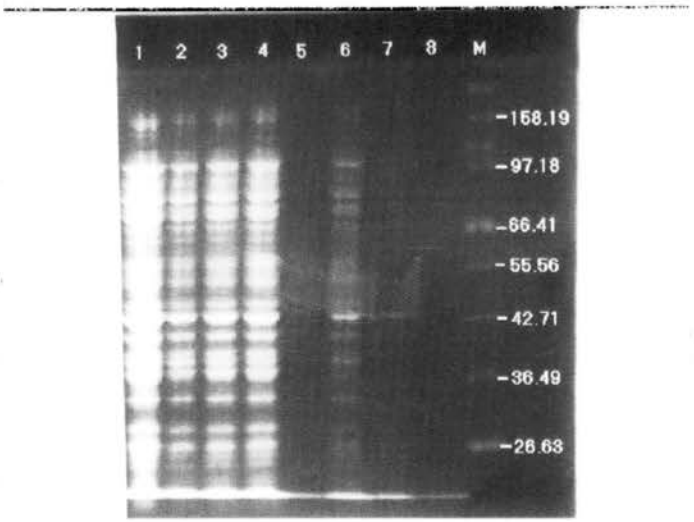


Figure 17. Induction of MBP-methylase fusion protein expression from pSJ136. The proteins were resolved on a 10% SDS-PAGE gel. Lane M, the molecular masses (in kDa). Lane 1, the uninduced sample. Lane 2, 3 and 4 are the one, two and three hour induced samples; lane 5, the osmotic shock fluid; lane 6, the crude extract; lane 7, the crude extract bound to the resin and lane 8, the proteins released by amylose from the resin.

induction (upon addition of IPTG), almost all of the cell's resources are converted to express the protein of interest, as the T7 RNA polymerase is very selective and active. As much as 50% of the total cellular protein would be comprised of the protein of interest within a few h of induction. The gene of interest can be maintained transcriptionally silent in the uninduced state. Recombinant plasmids that are not stable in the host can be stabilized by expressing the gene in hosts containing compatible plasmids that provide a small amount of T7 lysozyme that has been shown to be a natural inhibitor of the T7 RNA polymerase (139, 195). The two plasmids that express the T7 lysozyme are pLysS and pLysE. The expression of the lysozyme from pLysS is lower when compared to that of pLysE. The low-level of lysozyme produced by pLysS has little effect on the expression of the gene of interest following induction whereas the expression of the gene of interest is substantially decreased in the hosts carrying pLysE.

Cloning ORF6α into pET30b$^+$

The methylase gene was subcloned from pSJ136 as follows: pSJ136 was digested with *Ecl*136II and *Hin*dIII and the 1.4 kb fragment released was gel purified and ligated to pET30b$^+$ digested with *Eco*RV and *Hin*dIII. The recombinant plasmid, designated pSJ137, was subjected to restriction enzyme digestion to confirm the integrity of the plasmid. The nucleotide sequence of pSJ137 was obtained and, as expected, the methylase gene was in frame under the control of the T7 promoter.

Expression of the (ORF6α) Methylase

In order to optimize the over-expression of the methylase and select the host strain to be used, a pilot induction was done with three different hosts. pSJ137 was transformed into *E. coli* BL21(DE3) cells and also into BL21(DE3) host cells carrying either pLysS or pLysE. The strains were grown as previously described and induced with 1mM IPTG for three h and samples were taken at the end of every hour. The samples collected were electrophoresed on a SDS-PAGE and the expression of ORF6α was checked. As evident from the results, shown in Fig. 18, the expression was adequate neither in BL21(DE3) hosts nor the host carrying pLysE. Also, the level of expression increased and reached a maximum after 3 h of induction. So BL21(DE3) pLysS cells were chosen as the optimal host and 3 h of induction was taken to be the optimal time required. It was also observed that for reasons that are not clear at present, extended subculturing of the *E. coli* cells led to a very low level expression of the methylase. Glycerol was also found to lower the expression level. These problems were overcome by growing an isolated colony in the broth and using the broth culture for inoculating while doing the large-scale protein purification.

Protein purification from the inclusion bodies

The large-scale protein purification was done as previously described. The SDS-PAGE analysis of the fractions collected is shown in Fig. 19. It was observed that

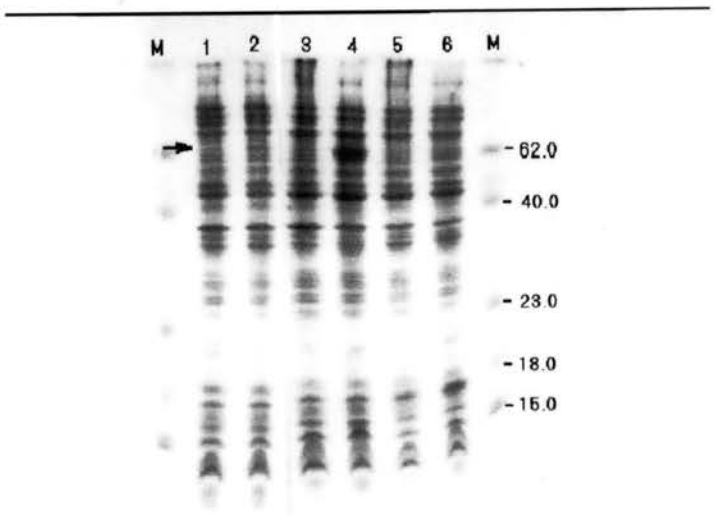


Figure 18. Determination of the optimal expression host for expressing the methylase fusion protein by SDS-PAGE analysis. The nominal molecular masses in kilodaltons are indicated (lane M). Lanes 1 and 2, lysates from BL21(DE3) hosts containing pSJ137; lanes 3 and 4, lysates from BL21(DE3) hosts containing pSJ137 and pLysS; lanes 5 and 6, lysates from BL21(DE3) hosts containing pSJ137 and pLysE. The uninduced lysates from the three different hosts are in lanes 1, 3 and 5. Lanes 2, 4 and 6 have the three hour induced lysates. The arrow shows the induced methylase.

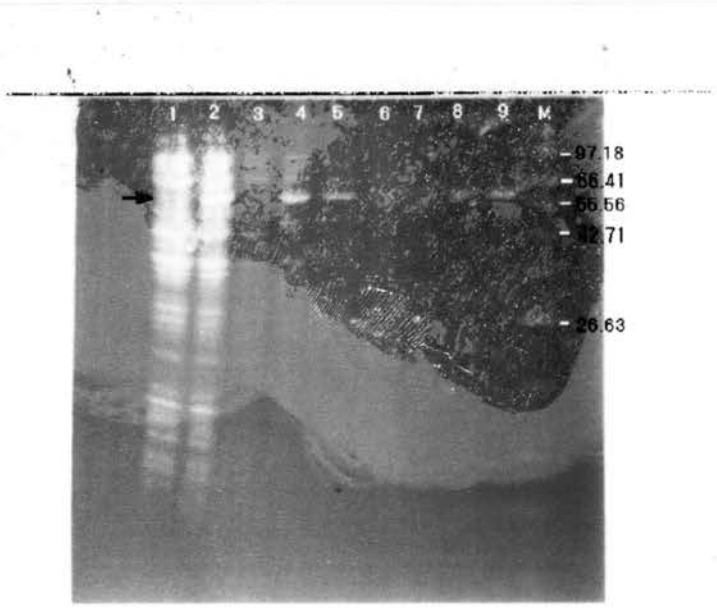


Figure 19. Large scale purification of the methylase fusion protein from the insoluble fraction. Lane M, the molecular marker and the masses indicated are in kDa. Lane 1, the uninduced lysate; 2, the three-hour induced lysate; 3, the soluble fraction; 4, the insoluble fraction; 5, the filtered insoluble fraction, 6, the column unbound fraction. Lanes 7, 8 and 9 are 8 μg, 20 μg and 40 μg, respectively of the pooled fractions containing the methylase.

about 6% of the total cellular proteins in the induced cells was made up of the methylase fusion protein product. However, most of the ORF6α fusion protein was insoluble. From 1 l of induced cells, approximately 4 mg of the ORF6α fusion protein carrying the His-Tag sequence was isolated using affinity column chromatography with charged His-Bind metal chelation resin under denaturing conditions in 6M urea. The purified protein fractions were pooled, renatured by step dialysis, and stored. The stored fusion protein was later used in methyltransferase assay along with the crude extract from the host strain harboring pSJ137. Surprisingly, the purified protein fraction was found to lack enzymatic activity (Table VI) although the total soluble protein extract of the cells carrying pSJ137 was capable of methylating DNA.

Cloning of ORF6β in an expression vector

ORF6β was amplified by PCR using ORF6β and JRRK1-6 primers. The template used was pSJ107 linearized with *Hind*III. The blunt ended fragment obtained was purified from the gel and ligated to *Eco*RV digested pET30a^{+}. The ligation was transformed into *E. coli* JM109 competent cells and the colonies obtained were used for mini plasmid DNA preparations. It was observed that in most cases there were deletions that spanned even the vector plasmids. As in the case of ORF6α cloning, XL1-Blue MRA cells were used as the host and transformants were replica-plated and used for colony hybridization experiments using PCR amplified ORF6β as the probe.

Table VI. Methylase assay using the fractions collected during the purification of the methylase from the inclusion bodies.

PROTEIN	cpm
Filtered Soluble Extract	354
***E. coli* XL1 Blue MRA (pSJ107)**	521
Purified methylase	11

Six microgram of the extract or 1 μg of the purified protein was added and the substrate DNA (chromosomal DNA of GM2163) used for the assay was 5 μg.

The probe failed to hybridize to any one of the colonies indicating that there were no colonies that contained the insert. The cloning attempts were repeated, but ORF6β could not be cloned in the expression vector.

Production of anti-ORF6α antibody

The methylase obtained from the insoluble fraction was used to generate polyclonal antibodies. The antibodies were raised in mice and the process was carried out by the Hybridoma Center of Oklahoma State University.

Optimization of the production of methylase

As most of the induced protein was present in the insoluble fraction, the optimal concentration of IPTG needed for adequate induction of the methylase and for the presence of maximal amount of the methylase in the soluble fraction was determined. Fig. 20 shows the SDS-PAGE analysis of the soluble fraction obtained from cells induced with varying amounts of IPTG for various lengths of time. The methylase assay was also performed using these fractions obtained (Table VII). It was evident that the activity of the soluble fraction increased as the concentration of IPTG is increased. The maximum activity was observed after 3 h induction with 1mM IPTG. Fig. 21 shows the western blot analysis of the soluble fractions obtained from cells induced with 0.5mM IPTG and 1mM IPTG for 1 h, 2 h and 3 h. The unpurified polyclonal antibodies raised

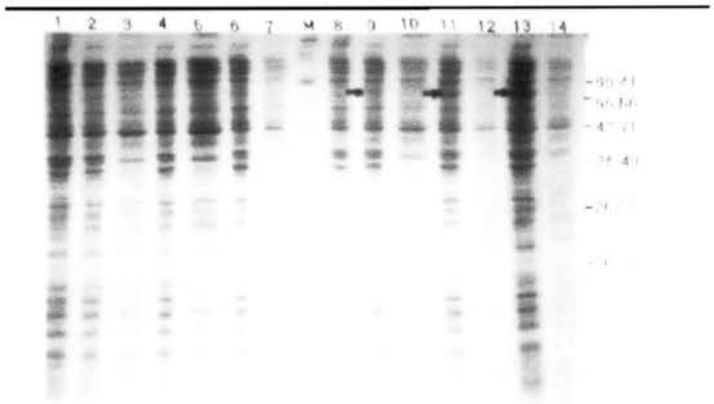


Figure 20. SDS-PAGE analysis of the soluble fraction obtained from pSJ137 carrying cells induced with varying amounts of IPTG. The molecular masses indicated (in kDa) correspond to the proteins present in lane M. Lanes 1 and 8, the uninduced lysates; lanes 1-7, lysates from BL21(DE3) containing pSJ137 and pLysS cells that were induced with 0.05mM IPTG; lanes 8-14, lysates from BL21(DE3) containing pSJ137 and pLysS cells that were induced with 1mM IPTG; lanes 3, 5, 7,10, 12 and 14, the soluble fractions obtained from different induced samples. Lanes 2, 3, 9 and 10, one hour induced samples; 4, 5, 11 and 12, two hour induced samples; 6, 7, 13 and 14, three hour induced samples. The arrows show the induced methylase.

Table VII. Methylase assay using the soluble fractions obtained from expression hosts carrying pSJ137.

IPTG USED FOR INDUCTION (mM)	HOURS OF INDUCTION	cpm
0.003	1	0
	2	0
	3	20
0.006	1	6
	2	41
	3	0
0.01	1	13
	2	0
	3	7
0.05	1	419
	2	458
	3	1456
1.0	1	1715
	2	2638
	3	2999
1.0 (pET30b)	3	0

Six microgram of the extract was used with 5 μg of the substrate DNA (GM2163 DNA)

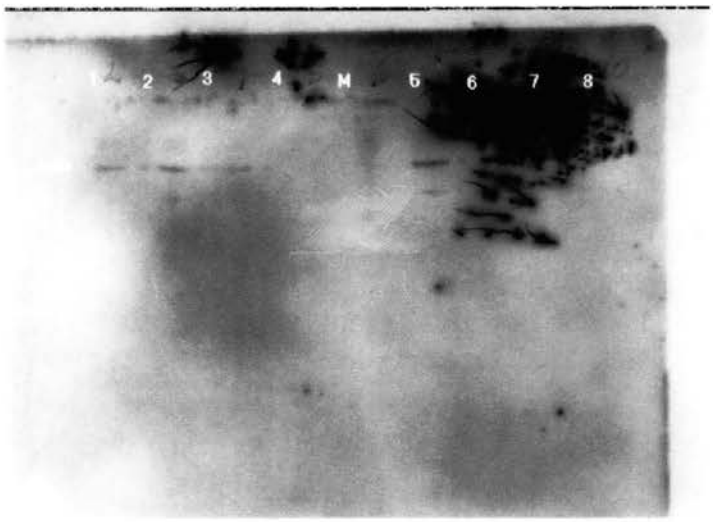


Figure 21. Western Blot analysis of the soluble fraction obtained from cells carrying pSJ137, induced with 0.05 mM and 1.0 mM IPTG. M, the prestained molecular weight standard. Lanes 1-4, the lysates from BL21(DE3) containing pSJ137 and pLysS cells that was induced with 0.05 mM IPTG; lanes 5-8, lysates from the same host cells but induced with 1 mM IPTG. Lanes 1 and 5, the uninduced sample; 2 and 6, soluble fraction from cells induced for 1 hour; 3 and 7, soluble fractions from cells induced for 2 hours; lanes 4 and 8, soluble fraction obtained after inducing the cells for 3 hours. The arrow shows the induced methylase.

against ORF6α cross-reacted with a number of othr proteins. To eleminate the non-specific antibodies, the polyclonal antibodies obtained was mixed with the crude lysate of over expressed *E. coli* host cells (BL21DE3 pLysS) containing the vector pET30b$^+$.

Purification of the methylase from the Soluble Fraction

As the purified protein obtained from the insoluble fraction was found to be non-functional, attempts were made to purify the protein from the soluble fraction. Using the optimized conditions, the soluble fraction was obtained from a 4 l culture and used for affinity chromatography using His-Bind metal chelation resin. The protein content of the fractions obtained was analyzed and a SDS-PAGE was performed (Fig. 22). The maximum amount of ORF6α protein was found to be in fractions 3, 4 and 5. These fractions were dialyzed overnight in a buffer containing 10mM Tris-HCl (pH 7.4), 25mM KCl, 1mM DTT, 10μM AEBSF and 50% glycerol. The next day a significant amount of precipitate was observed in the dialysis bags. The supernatant fluid was taken and stored. As evident from the SDS-PAGE analysis, most of the methylase was lost during the precipitation. Nevertheless, the fractions obtained before and after dialysis were used in the methylase assay (Table VIII). Fraction 3 showed higher activity than fractions 4 and 5 after dialysis. This decrease in the activity of fraction 4 and 5 was also evident from the SDS-PAGE gel (Fig. 22) as there was only a very faint band corresponding to the methylase. The increase in the methylase activity of fraction 3 after dialysis could be due to the removal of NaCl and other substances (such as Imidazole) present in the

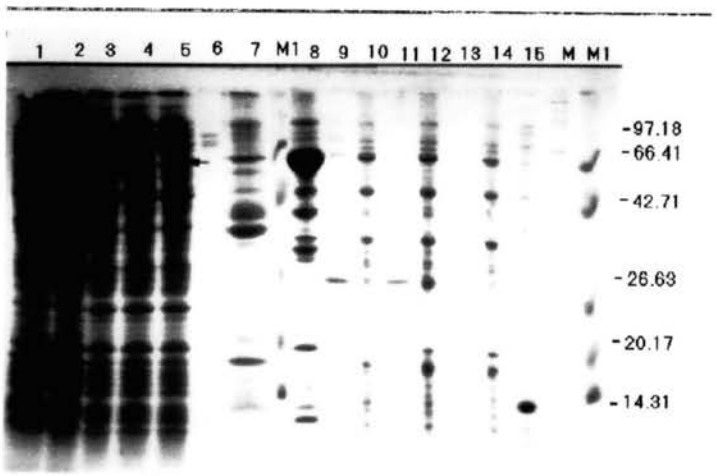


Figure 22. SDS-PAGE analysis of the fractions collected during the purification of the methylase from the soluble fraction. M is the molecular weight standard and the weights indicated are in kiloDaltons. Lane 1, the uninduced sample; 2, the 3 hour induced sample; 3, the insoluble fraction; 4, the unfiltered soluble fraction; 5, the filtered soluble fraction; 6, the fraction unbound to the column; 7, the column wash; 8, the purified but inactive methylase. Lanes 9, 11, 13, the fractions 3, 4 and 5 respectively after dialysis; lanes 10, 12, and 14, the precipitates obtained from fractions 3, 4 and 5 respectively during dialysis. Lane 15, the extract from XL1 Blue MRA cells carrying pSJ107.

Table VIII. Methylase assay using the fractions collected during the purification of the methylase from the soluble fraction.

PROTEIN	cpm
Filtered Soluble Extract	7466
Unbound	2836
Fraction #3 before dialysis	4439
Fraction #3 after dialysis	9813
Fraction #4 before dialysis	3984
Fraction #4 after dialysis	446
Fraction #5 before dialysis	1569
Fraction #5 after dialysis	673
***E. coli* XL1 Blue MRA (pSJ107)**	798
Inactive methylase	4

Six microgram of the extract or 1 μg of the purified fraction was used for the assay with 5 μg of the substrate (GM2163) DNA. Inactive methylase refers to the purified methylase obtained from the insoluble fraction. The fractions obtained were eluted from the column using the 1x elute buffer containing 1 M imidazole.

elution buffer. The protein was purified again from a 2 l culture, using His-bind affinity chromatography with the intention of running an ion exchange chromatography using a DEAE-Sepharose column. A Bradford assay (19) was run on the fractions collected from the affinity column. Some of the fractions collected were run on SDS-PAGE (Fig. 23) and were also used to check for the methylase activity (Table IX). Fraction 3, which showed the maximal methylase activity, was dialyzed overnight in DEAE starting buffer containing 20 mM Tris-HCl (pH 7.4), 5 mM NaCl, 1 mM EDTA, 1 mM DTT and 1 μM AEBSF. The next day even in this case a significant amount of precipitate was observed. Unlike the previous cases, a Bradford assays revealed that there was no protein present in the supernatant fluid. The problem of precipitation was solved by dialyzing the protein fractions in a buffer containing 50 mM Tris-HCl (pH 7.4), 200 mM KCl, 10 mM EDTA, 1 mM DTT and 50% glycerol. The fusion protein was purified again from 2 l culture using affinity chromatography and the bound protein was eluted with elution buffer containing 500 mM NaCl, 20 mM Tris-HCl (pH 7.9) and varying amounts of imidazole, 200 mM, 400 mM, or 1 M. A Bradford assay was done on fractions collected and some of the fractions were run on SDS-PAGE (Fig. 24). The same fractions were also used for methylase assays (Table X) and surprisingly the methylase activity was observed in fractions eluted with all the different concentrations of imidazole. Two fractions that showed the highest incorporation were pooled and concentrated using an Amicon column. This fraction was given to the DNA/Protein Recombinant DNA/Protein Resource facility of Oklahoma State University for purifying the protein by

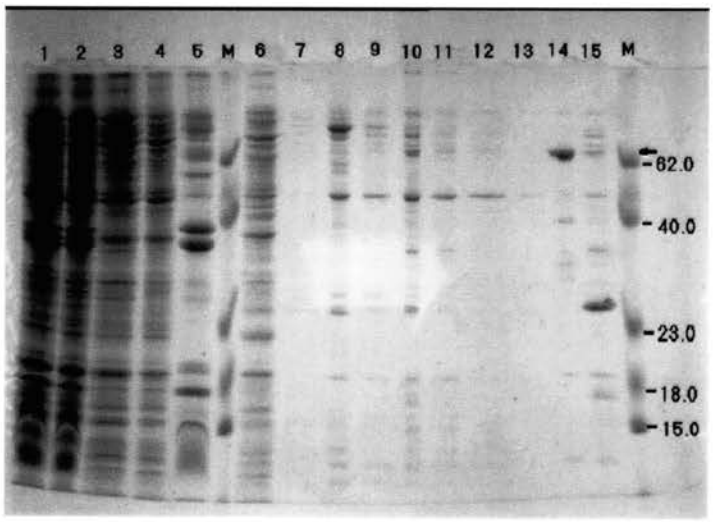


Figure 23. SDS-PAGE analysis of the different fractions collected during the purification of the methylase from the soluble fraction for the second time.

M, the molecular weight standard and the weights indicated are in kilodaltons. Lane 1, the uninduced sample; 2, the 3 hour induced sample; 3, the unfiltered soluble fraction; 4, the filtered soluble fraction; 5, the insoluble fraction; 6, the fraction unbound to the column; 7, the binding buffer wash; 8, the column wash. Lanes 9-13, the fractions #2, 3, 4, 5 and 8 collected respectively; lane 14, the purified but inactive methylase; lane 15, the previously purified active methylase. Arrow points to the methylase protein.

Table IX. Methylase assay using the fractions collected during the second time.

PROTEIN	cpm
Unfiltered Soluble Extract	372
Filtered Soluble Extract	426
Unbound Extract	0
Wash	822
Fraction #2	257
Fraction #3	1216
Fraction #4	166
Fraction #5	81
Fraction #8	125
Active methylase	14356
Inactive methylase	0

Six microgram of the crude protein extract or 1 μg of the purified fraction was used for the assay with 5 μg of the substrate (GM2163) DNA. Active methylase refers to the protein purified from the soluble fraction (fraction 3 of the first isolation) that was given for HPLC analysis and the inactive methylase refers to the protein purified from the soluble fraction. The fractions were collected using 1 x elute buffer containing 1 M imidazole.

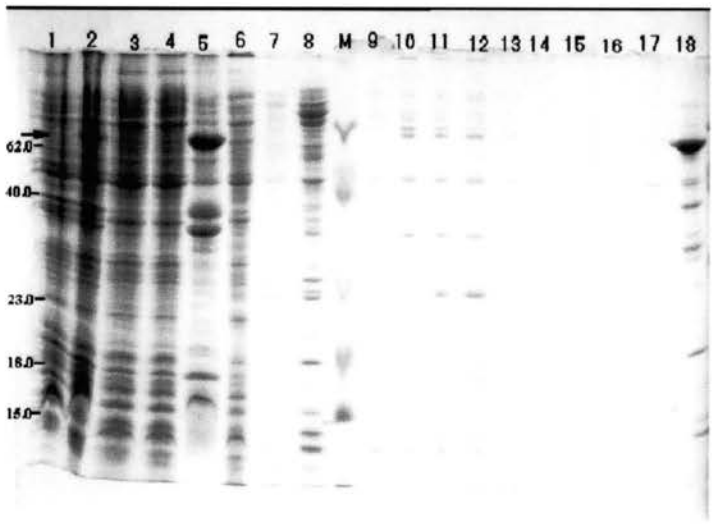


Figure 24. SDS-PAGE analysis of the fractions collected during the methylase purification with step gradient. M, the molecular weight standard and the weights indicated are in kiloDaltons. Lane 1,the uninduced sample; 2,the 3 hour induced sample; 3, the unfiltered soluble fraction; 4, the filtered soluble fraction; 5, the insoluble fraction; 6, the fraction unbound to the column; 7, the binding buffer wash; 8, the column wash. Lanes 9-12, the fractions collected using 200 mM Imidazole; lanes 13-15, the fractions collected using 400 mM Imidazole; lanes 16 and 17, the fractions obtained using 1 M Imidazole; lane 18 is the purified but inactive methylase. Arrow points to the methylase protein.

Table X. Methylase assay using the fractions obtained during the step gradient purification.

IMIDAZOLE CONCENTRATION	PROTEIN	cpm
200mM	Fraction #2	6160
200mM	Fraction #6	14425
200mM	Fraction #7	11364
200mM	Fraction #8	12388
400mM	Fraction #1	3912
400mM	Fraction #3	11672
400mM	Fraction #4	4344
400mM	Fraction #8	5998
1000mM	Fraction #5	9600
1000mM	Fraction #7	11159
---	Active methylase	18107
---	Inactive methylase	0

One microgram of the purified fraction was used for the assay with 5 μg of the substrate (GM2163) DNA. Active methylase refers to the purified farction (fraction 3 of the first isolation) obtained from the soluble fraction that was given for the HPLC analysis. Inactive methylase refers to the protien purified from the insoluble faction.

gel-filtration uisng HPLC. HPLC analysis using the Biorad Superdex-75 column was done and the proteins were eluted using the same buffer that they were stored in but without the 50% glycerol. The fractions were collected and the absorbance was measured simultaneously (Fig. 25). The profile shows a peak at fraction 6 and by comparing this peak with the standard, the size of the protein in the peak was estimated to be above 100 kDa, which is much larger than that expected for ORF6α. The fractions collected were checked for their protein content and used in methylase assays. It was found that there was no methylase activity (Table XI). Some of the HPLC fractions collected were also electrophoresed on a SDS-PAGE (Fig. 26). Analysis of the SDS-PAGE gels showed a very low concentration, of a protein band corresponding to that of ORF6α in some fractions. The presence of a peak of about 100 kDa in the HPLC was assumed to be due to hydrophobic interactions which resulted in the aggregation of the ORF6α so the HPLC analysis was performed again using a different fraction of purified protein. The elution buffer used was 50 mM Tris-HCl (pH 7.4), 10 mM EDTA, 0.5 mM DTT, 0.1% Triton-X-100 and 0.05% Sarkosyl. The elution buffer used was the same as in the previous HPLC analysis except Triton and Sarkosyl were incorporated to overcome the hydrophobic interactions. The analysis revealed a peak corresponding to a protein of over 100 kDa and a smaller peak corresponding to a protein of about 70 kDa was also observed (Fig. 27). The fractions were checked for activity and as before no activity was observed (Table XII). Although the purified protein fraction used for HPLC was active, for reasons not known at present the HPLC fractions collected were not.

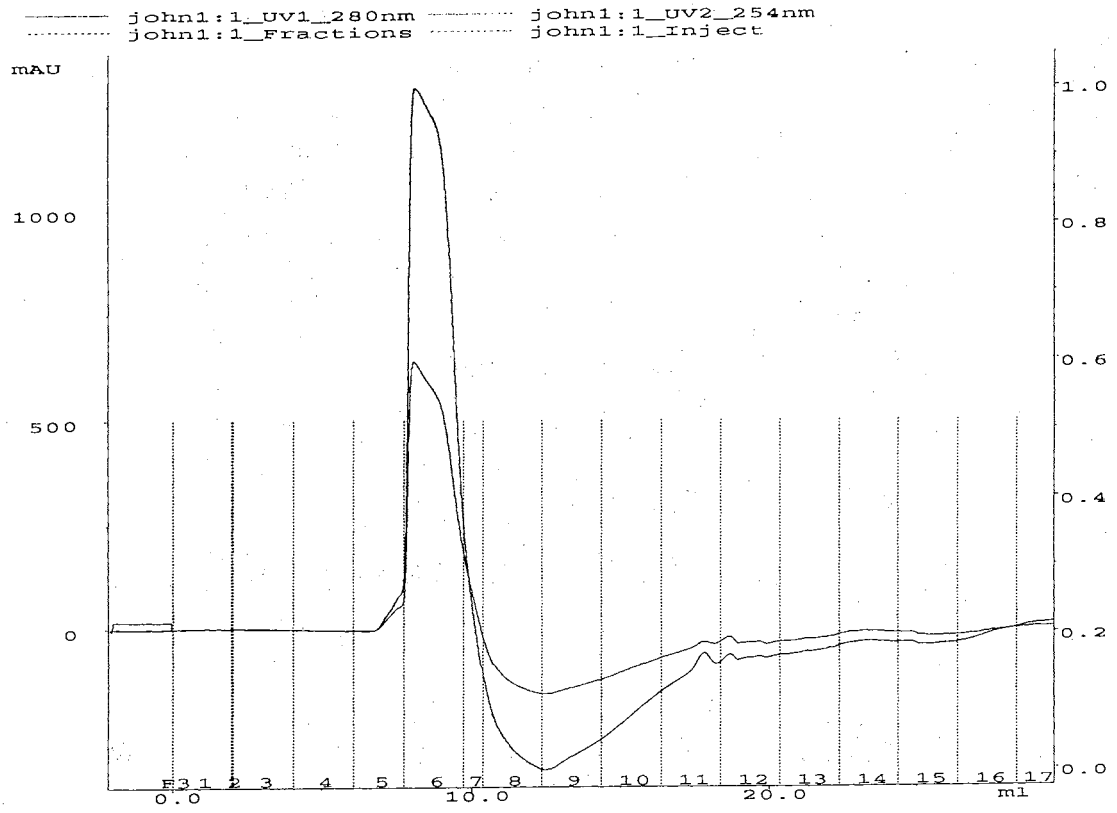


Figure 25. UV absorbance during the purification process by HPLC. The absorbance at 280nm and 254nm were measured and recorded along with the fraction numbers.

Table XI. Methylase assay using the HPLC fractions.

PROTEIN	cpm
Purified Active methylase	451
Fraction 5 of HPLC	0
Fraction 6 of HPLC	0
Fraction 7 of HPLC	0
Fraction 8 of HPLC	26
Purified Inactive methylase	0

One tenth microgram of the purified methylase or fraction was used with 5 μg of the substrate DNA. The purified active methylase was given for HPLC analysis. Active methylase refers to the purified fraction obtained from the soluble fraction that was given for the HPLC analysis. Inactive methylase refers to the protein purified from the insoluble faction.

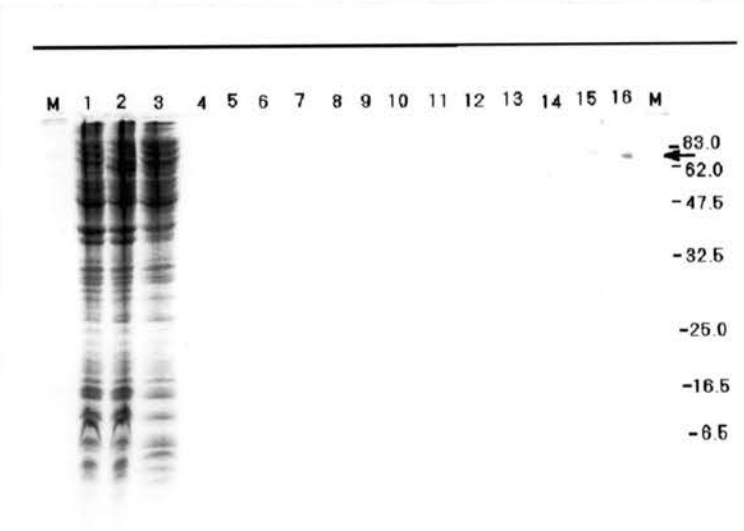


Figure 26. SDS-PAGE analysis of some of the HPLC fractions. M is the Molecular masses (in kDa). Lane 1, the uninduced sample; 2, the 3 hour induced sample; 3, the filtered soluble fraction. Lanes 4-14, the different fractions (#5, 6, 7, 8, 9, 10, 11, 12, 14, 15, and 17) collected during HPLC; lane 15, the methylase that was given for HPLC and lane 16, the purified but inactive methylase. Arrow points the methylase.

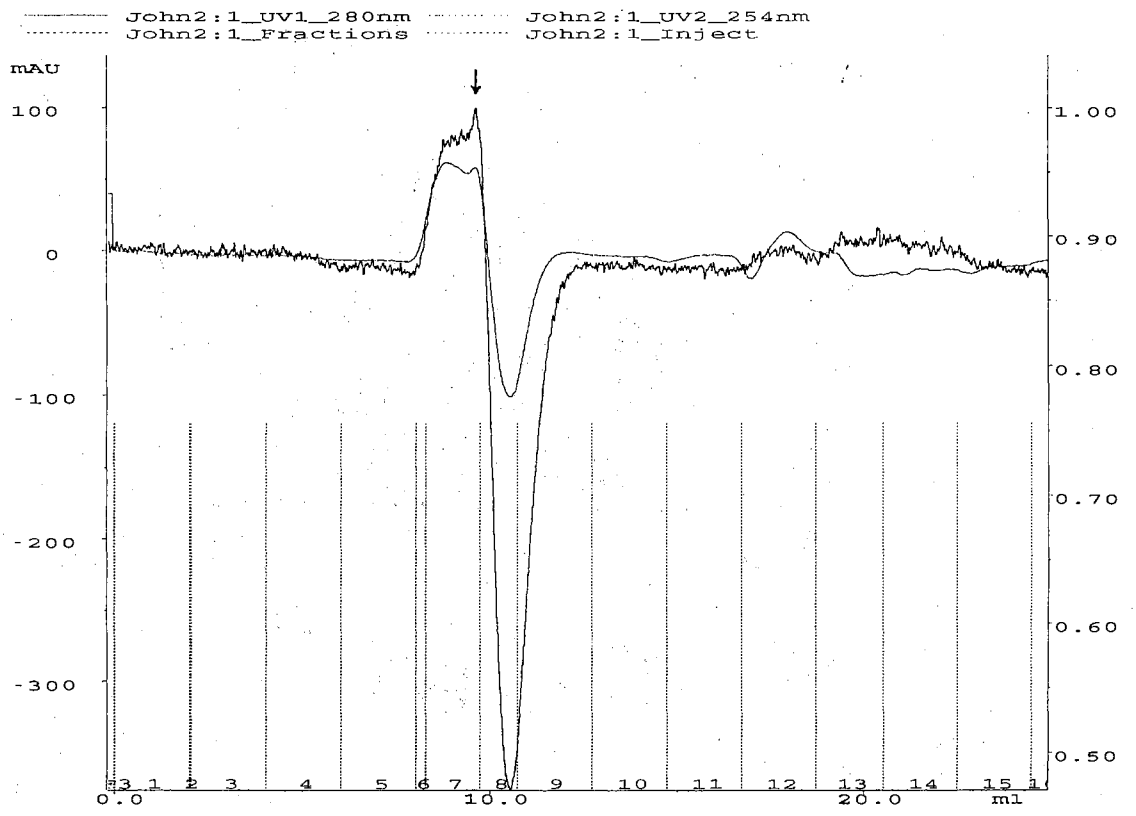


Figure 27. UV absorbance of the HPLC fractions collected using an elution buffer containing 0.1% Triton and 0.05% Sarkosyl. UV absorbance at 280nm and 254nm were measured and recorded. The graph also shows the fractions numbers and the arrow points to the new peak that was observed. At 280nm, for reasons not known at present a significant level of noise was observed.

Table XII. Methylase assay using the HPLC fractions collected using the buffer with 0.1% Triton and 0.05% Sarkosyl.

PROTEIN	cpm
Purified Active methylase	1434
Fraction 5 of HPLC	0
Fraction 6 of HPLC	0
Fraction 7 of HPLC	0
Fraction 8 of HPLC	0
Fraction 9 of HPLC	0
Purified Inactive methylase	0

Two tenths microgram of the purified methylase or fraction was used with 5 μg of the substrate DNA. The active methylase purifed from the soluble fraction used was given for HPLC analysis. The inactive methylase was obtained from the insoluble fraction.

Site of Methylation

Using one of the purified fractions that was active, pBR322 was methylated *in vitro* using non-radioactive S-adenosyl methionine. After overnight incubation at 37°C, the plasmid was cleaned and digested with several different restriction endonucleases (Fig. 28). *Bst*NI is an isoschizomer of *Eco*RII but it restricts DNA irrespective of the methylation pattern. Comparison of the *Bst*NI and *Eco*RII digests revealed that the *Eco*RII sites were protected by methylation. As evident, *Hae*III sites and the *Sau*3AI sites present in pBR322 were also protected by the methylation. Thus, the methylase methylates the cytosines in the *Eco*RII sites, the *Hae*III sites and the *Sau*3AI sites. Analysis of the restriction digestion suggests that the inner cytosine in the *Eco*RII site (CCWGG) is methylated. Identification of the target cytosine in the other target sequences awaits further study.

In vivo Mutagenesis of ORF6

The *E. coli* plasmid, pVA891 carrying a streptococcal erythromycin resistance determinant that is expressed in pneumococci when the plasmid is inserted into the chromosome, was used as a marker DNA segment for insertion mutagenesis. The strategy used for the site-specific mutagenesis of ORF6 is outlined in Fig. 29 and is similar to the one previously reported (97). Briefly, pSJ107 was linearized at the unique *Kpn*I site within ORF6, treated with the Klenow fragment to make the ends blunt ended

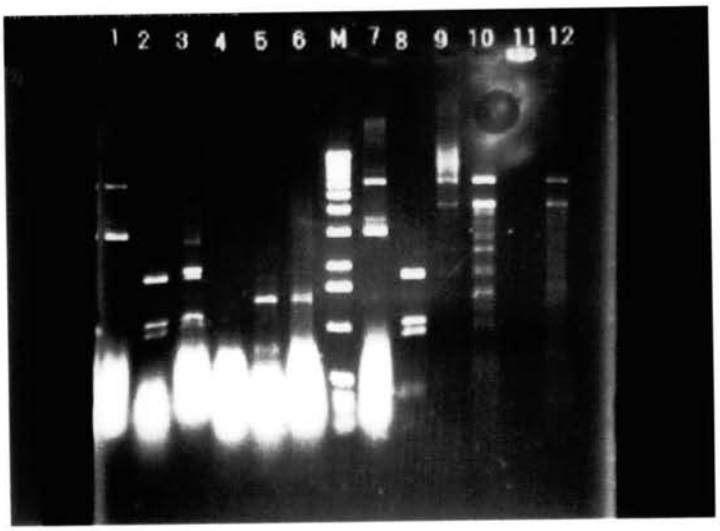
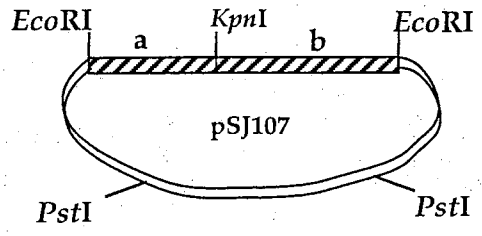
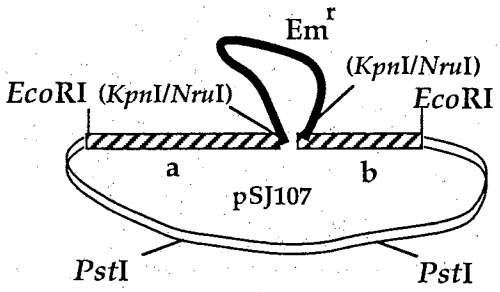


Figure 28. Agarose gel picture of the restriction digest of methylated pBR322. The lane marked (M) contains the 1 kb ladder and the sizes are in kilobases. pBR322 was methylated *in vitro* and subjected to restriction digestion with *Bst*NI, *Eco*RII, *Hae*III, *Mbo*I and *Sau*3AI (lanes 8-12 respectively). Unmethylated pBR322 was digested with the same enzymes (lanes 2-6 respectively) as a control. Lanes 1 and 7, contain uncut, unmethylated and methylated pBR322 DNA respectively.

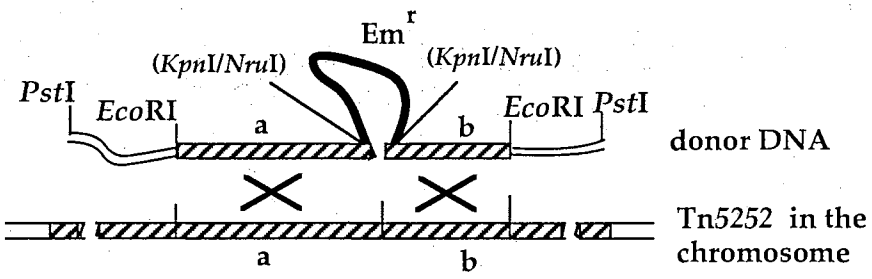
A - Clone the 3.27 kb *Eco*RI fragment of the transposon into pLS1



B - Insert the drug resistance marker into the *Kpn*I site of the passenger DNA.



C - Linearize the vector at the *Pst*I, and use it to transform SP1000 (Tn5252)



D - The resulting insertion mutant (Tn5252 ::Em^r)

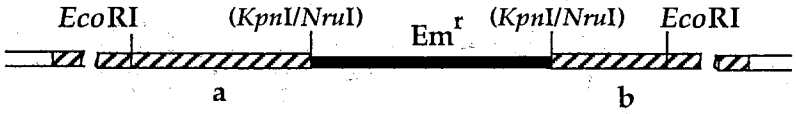


Figure 29. Strategy used for insertion mutagenesis of ORF6.

and ligated to pVA891 linearized with *Nru*I. Cleaving the pLS1 portion in pSJ107 following digestion with *Pst*I linearized the ligated molecules. The digested DNA was introduced by transformation into competent SP1000 cells carrying Tn*5252*. In this homology-dependent recombination event, the heterologous pVA891 was expected to be inserted due to the flanking homology provided by the ORF6 DNA. Chromosomal DNA from one of the Em^r transformants, SP1262, was used in transformation experiments.

Pneumococcal Transformation to check for the Em insertion

To check if the pVA891 inserted into the transposon, the chromosomal DNA from SP1262 was used as donor in filter-mating experiments. The control pneumococcal chromosomal DNAs used in this experiment were DP1617 and SP1255. The recipients used were Rx1 (the wild-type *S. pneumoniae*) and DP1322 (pneumococcus carrying the parent element, Tn*5253*). Before transformation, a sample of each DNA was sheared using a 21 1/2 gauge needle. Both of the recipients were transformed with each one of the sheared, as well as, the unsheared DNAs. The cells were then plated by the overlay method and transformants were selected either with streptomycin or with erythromycin. The results of the experiment are presented in Table XIII. If the insertion had taken place in the transposon, the frequency of Em^r transformants of DP1322 would be greater than that of Rx1 as Em^r transformants would

Table XIII. Transformation of Rx1 and DP1322 to check if the Em insertion is in the transposon.

DONOR DNA	MARKER	Rx1 TRANSFORMANTS /ml	DP1322 TRANSFORMANTS /ml
UNSHEARED	*str*	52000	42000
DP1617	Em	0	0
SHEARED	*str*	62000	49000
	Em	0	0
UNSHEARED	*str*	94000	53000
SP1255	Em	0	1260
SHEARED	*str*	79000	24900
	Em	10	360
UNSHEARED	*str*	19100	18500
SP1262	Em	0	1150
SHEARED	*str*	10600	11800
	Em	10	620

1 μg /ml of each DNA was used to transform competent Rx1 and DP1322 cells.

arise due to flanking homology provided by the transposon DNA. This high frequency of transformation of DP1322 was expected to decrease due to shearing of the donor DNA, as shearing decreases the flanking homology (because of the decrese in the size of the DNA). DP1617 was not expected to give rise to any Em^r transformants, as there is no erthromycin gene in its DNA. From the results, it is evident that the number of Em^r transformants of DP1322 was atleast three orders of magnitude greater than the number obtained in Rx1. This confirmed that the insertion of pVA891 had taken place in the transposon.

Southern Hybridization

To check if the intended mutation had take place, the chromosomal DNA of SP1262 was digested with *Eco*RI and *Hin*dIII, separated by electrophoresis on an agarose gel, and transferred to nylon membrane. Chromosomal DNAs from Rx1 and SP1000, digested with the same enzymes served as controls. The samples were probed with radiolabled pSJ107 DNA. The membrane was washed and exposed to X-ray film. Fig. 30 shows the restriction maps of the 3.27 kb *Eco*RI fragment with pVA891 inserted into the unique *Kpn*I site in either one of the two possible orientations. The autoradiogram is shown in Fig. 31. As expected, the probe reacted with a 2.1kb (target region) and a 3.27 kb (right junction fragment) *Eco*RI fragment of Rx1 and SP1000 DNAs, respectively. Depending on the orientation of the insertion of pVA891 within ORF6, the probe was expected to hybridize either with 6.43 and 2.75 kb or with 5.13 kb

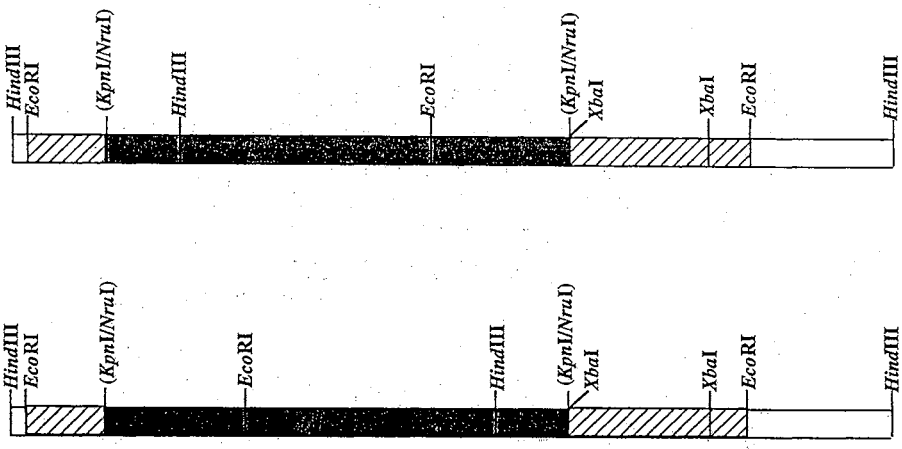


Figure 30. Restriction map of SP1262, the Em^r insertion mutant. The two possible orientations of the pVA891 in the *Kpn*I site of ORF6 are shown. The *Eco*RI fragments that would be obtained from the DNA represented by the upper figure would be 5.13 and 4.05 kb, the sizes from the lower one would be 2.75 and 6.43 kb. Crosshatched segment represents the3.27 kb *Eco*RI fragment of interest. Solid box indicates pVA891, the heterologous DNA used for insertion. Relevant restriction sites are shown.

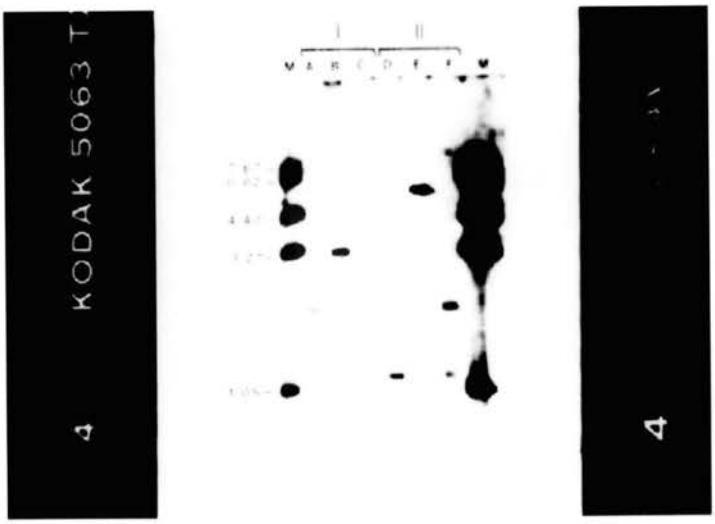


Figure 31. Autoradiogram showing the insertion of heterologous DNA within the ORF6. Chromosomal DNA from an Em^r transformant, SP1262, was isolated, digested with *Eco*RI or *Hin*dIII, elctrophoresed on a 0.8% agarose gel, blotted to a nylon membrane, and probed with radioactively labeled pSJ107. I, *Eco*RI digests; II, *Hin*dIII digests; Lanes A and D, DNA from Rx1; B and E, DNA from SP1000 (Rx1::Tn*5252*); C and F, DNA from SP1262. and M, molecular weight markers. Nominal molecular masses are indicated.

and 4.05 kb *Eco*RI fragments from SP1262. As evident from the gel, the probe reacted with DNA fragments of the latter size indicating the orientation (see Fig.30). Also, the hybridization of pSJ107 with 9.15 and 2.13 kb *Hind*III fragments (laneF) confirmed the orientation as well as the integrity of the inserted pVA891.

Transformation of DP1002

In order to check the effect of the presence of multiple copies of the methylase in the recipient during conjugal transfer, a pneumococcal strain carying pSJ107 had to be created. Competent DP1002 cells were transformed with pSJ107 as described previously. The overlay method was used to select the transformants and the selection used was tetracycline (plasmid marker). The plasmid was isolated from one of the Tc^r nov^r transformant and subjected to restriction enzyme digestion, to make sure that the plasmid was intact.

Conjugal Transfer Properties of the Insertion Mutant

To determine the role of ORF6 in the conjugal transfer of Tn*5252*, SP1262 was used as a donor in filter-mating experiments. SP1254 carrying pVA891 inserted at a locus within the element that does not impair its transfer functions and provided the control (97). The filter-matings were performed as described previously (7). After mating, the transconjugants were selected on CAT agar containing the recipient marker (novobiocin in the case of DP1002) and donor marker (erythromycin). The Em^r

transconjugants were screened for the unselected markers, chloramphenicol (transposon marker), and streptomycin (chromosomal marker of donor strains). Concentrations of the selective drug used are listed in Table III. The transfer property of the ORF6-insertion mutant in the presence of multiple copies of the methyltransferase in the recipient is listed in Table XIV. Only a two-to-three-fold reduction in the transfer frequency of Tn*5252* from SP1262 was observed, as compared to SP1254. The presence of multiple copies of the methylase (pSJ107) in the recipient also decreased the transfer frequency by about two-fold. Thus, the insertion mutation in ORF6 does not affect the conjugal transfer of the transposon.

Construction of pSJ119

In order to create a deletion mutation in ORF6, pSJ107 was digested with *Xba*I, as there were two sites in the passenger DNA and none on the vector (pLS1) portion. The large fragment (5.88 kb) was electroeluted, self ligated, and used to transform XL-1 Blue MRA. The plasmid was isolated from a Tc^r transformant and was subjected to restriction digestion to confirm the loss of the 1.8-kb *Xba*I fragment. The recombinant plasmid obtained was designated pSJ119. This mutation was expected to delete the amino-terminus of ORF6 and ORF41.

Table XIV. Transfer property of SP1262 in the presence of multiple copies of the methylase in the recipient.

DONOR	RECIPENT	PLASMID IN RECIPIENT	FREQUENCY (transconjugants/donor x 10^6)
SP1254	DP1002	None	598 ± 49
	SP1322	pSJ107	232 ± 15
SP1262	DP1002	None	175 ±13
	SP1322	pSJ107	71± 8

The results are from 2 independent experiments representing the range. SP1254 carries Tn*5252* with pVA891 inserted at a locus that is not involved in transfer and is str^r. SP1262 is Tn*5252* with pVA891 inserted within the ORF6 and is str^r. SP1322 is DP1002 containing pSJ107. DP1002 is nov^r, *S. pneumoniae*.

Construction of SP1273

For reasons not known at present, the presence of novobiocin resistance in the recipient during conjugation seemed to interfere with the conjugal transfer process. To avoid this problem, all the donor markers were changed from streptomycin to novobiocin by conjugation. After the filter-mating, one true transconjugant (*nov*^{r}, Cm^r, Em^r, *str*^{s}) from SP1254 x DP1002 as well as SP1262 x DP1002 matings was chosen and these were designated SP1265 and SP1273. Further conjugations were done with these strains as donors and DP1004 (*str*^{r}) as recipient.

Transformation of SP1273

To check the effect of the presence of pSJ107 or its derivatives in the donor cells, SP1273 was made competent using competent Rx1 supernatant fluid and transformed with pLS1, pSJ107, pSJ117 or pSJ119 as described previously. The transformants were selected with tetracycline and the presence of unselected markers (novobiocin, chloramphenicol, and erythromycin) was checked by replica plating. One transformant from each transformation experiment was chosen for further study. The plasmid DNA was isolated from each of the selected transformant to check for the presence and integrity of the plasmid. These strains were designated SP1296 (SP1273 ::pLS1), SP1297 (SP1273::pSJ107), SP1298 (SP1273::pSJ117), and SP1299 (SP1273::pSJ119) and were stored for future use.

Conjugation of the methylase insertion mutant with *S. pneumoniae*

To check if the insertion mutation in the methylase (ORF6) affected the transfer of the transposon, the above mentioned strains were used as donors in conjugation experiments with DP1004 as the recipient. After mating, the transconjugants were selected on CAT agar containing streptomycin and erythromycin. The Em^r transconjugants were then replica-plated for unselected markers (transposon marker and donor marker). The results are shown in Table XV. Comparison of the frequency of transfer of SP1265 and SP1273 indicated that the transfer of the transposon was not affected by the ORF6 insertion mutation, although it is reduced by about two-fold. However, the presence of pLS1 in the donor reduces the transfer frequency by about four-to six-fold.

Conjugal transfer of the methylase insertion mutant transposon to *S. pyogenes*

To check if the transfer of the transposon to a heterologous host is affected by the insertion mutation in the methylase gene, conjugation experiments were carried out with the *S. pyogenes* as recipient. The donor strains were created by transforming SP1262 with pLS1 to generate SP1292, pSJ107 to generate SP1293, pSJ117 to generate SP1294, and pSJ119 to generate SP1295. These newly created strains were checked for the presence of all the markers and the plasmids from each one was isolated to check for their integrity. These strains were used as donors in the filter-mating

Table XV. Effect of methylase mutation on the conjugal transfer of the transposon to *S. pneumoniae*.

DONOR	PLASMID IN DONOR	RECIPIENT	TRANSCONJUGANTS /DONOR x 10^7
SP1265	None	DP1004	200 ± 14
SP1273	None	DP1004	140 ± 12
SP1296	pLS1	DP1004	30 ± 5
SP1297	pSJ107	DP1004	36 ± 6
SP1298	pSJ117	DP1004	53 ± 7
SP1299	pSJ119	DP1004	60 ± 8

The results are from two independent experiments representing the range. SP1265 carries Tn*5252* with pVA891 inserted at a locus that is not involved in transfer and is nov^r. SP1273 is Tn*5252* with pVA891 inserted within the ORF6 and is nov^r. SP1296, SP1297, SP1298, and SP1299 are the same as SP1273 but carry pLS1, pSJ107, pSJ117 or pSJ119 respectively.

experiments. After mating, the transconjugants were selected on CAT agar plates containing optochin (recipient marker) and erythromycin (donor marker). The results obtained are listed in Table XVI. As in the case of pneumococcal recipients, the transfer of the transposon from the methylase mutant is reduced two-fold compared to that of SP1254. The presence of pLS1 or its derivatives in the SP1262 donors were not found to abolish the transfer frequency of the element to pneumococcal recipients. However, for reasons that are not clear at present, the presence of pLS1 or any one of the recombinant plasmids in the donor dramatically reduces the transfer of the transposon to *S. pyogenes*.

Conjugal Transfer of the methylase insertion mutant Tn*5252* to *S. pneumoniae* (*Dpn*II$^+$)

Guild, *et al.* (64) reported that the transfer of the transposon (Tn*5252*) to a pneumococcal strain that is proficient in *Dpn*II restriction modification system was not affected whereas simultaneous transfer of the conjugative plasmid, pIP501, from the same donor was inhibited by at least a 1000 fold. To check if the methylase present in the transposon played any role in protecting the transposon during conjugation, the insertion mutant created was used as donor in filter mating experiments. The results, listed in Table XVII, indicated that the transfer was not affected by the insertion mutation or by the presence of pLS1. However, the presence of multiple copies of the 3.27 kb fragment reduces the transfer frequency by 2-3 fold compared to SP1265.

Table XVI. Effect of the methylase insertion mutation on the conjugal transfer of the transposon to *S. pyogenes*.

DONOR	PLASMID IN DONOR	RECIPIENT	TRANSCONJUGANTS /DONOR x 10^8
SP1254	None	*S. pyogenes*	500 ± 22
SP1262	None	*S. pyogenes*	270 ±16
SP1292	pLS1	*S. pyogenes*	<1
SP1293	pSJ107	*S. pyogenes*	<1
SP1294	pSJ117	*S. pyogenes*	<1
SP1295	pSJ119	*S. pyogenes*	<1

The results are from two independent experiments representing the range. SP1254 carries Tn*5252* with pVA891 inserted at a locus that is not involved in transfer and is str^r. SP1262 is Tn*5252* with pVA891 inserted within the ORF6 and is str^r. SP1292, SP1293, SP1294, and SP1295 are the same as SP1262 but carry pLS1, pSJ107, pSJ117 or pSJ119 respectively. The recipient used was *S. pyogenes* ATCC 19615.

Table XVII. Conjugal transfer properties of the methylase insertion mutant to *S. pneumoniae DpnII*$^+$ host.

DONOR	PLASMID IN DONOR	RECIPIENT	TRANSCONJUGANTS /DONOR x 10^8
SP1265	None	DP3111	170 ± 13
SP1273	None	DP3111	210 ± 14
SP1296	pLS1	DP3111	300 ± 17
SP1297	pSJ107	DP3111	54 ± 7

The results are from two independent experiments representing the range. SP1265 carries Tn*5252* with pVA891 inserted at a locus that is not involved in transfer and is *nov*^{r}. SP1273 is Tn*5252* with pVA891 inserted within the ORF6 and is *nov*^{r}. SP1292 and SP1293 are the same as SP1273 but carry pLS1 and pSJ107 respectively. The recipient (DP3111) used was *S. pneumoniae* that is *Dpn*II$^+$ and *str*^{r}.

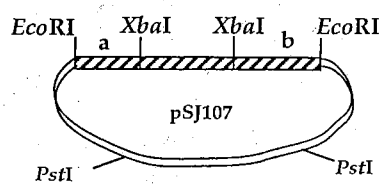
Construction of pSJ144

In order to generate a plasmid with a deletion in ORF6, pSJ119 was digested with *Xba*I and ligated to *Xba*I digested pVA891 to generate pSJ144. This plasmid was established in *E. coli* C600 host cells and the restriction map was confirmed by digesting with various restriction enzymes. The establishment of pSJ144 without any deletion in C600 itself indicated that the methylase was not functional.

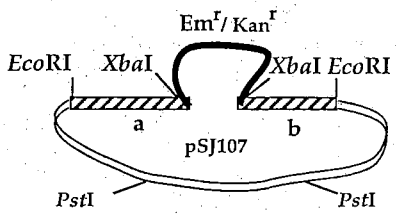
In vivo deletion mutation of the methylase to generate SP1213

As evident from Table V, the insertion mutation (in pSJ117) of the methylase showed an elevated methylase activity. This indicated that the truncated methylase was still active. To mutate the methylase, an alternate strategy (shown in Fig. 32) was used. pSJ144 was linearized by cleaving the pLS1 portion of pSJ119 with *Pst*I. The digested DNA was introduced by transformation into competent SP1000 cells. In this homology dependent recombination event, pVA891 was expected to be inserted due to flanking homology provided by the 3.27-kb DNA. Transformants obtained were replica-plated for unselected markers as before and seven true transformants were chosen for further study. The chromosomal DNA from these seven transformants were prepared and used as donors in transformation experiments with Rx1 and DP1322 cells as recipients. Table XVIII shows the results obtained from this experiment and it is evident that the insertion had taken place in the transposon in at least in five of the six clones chosen.

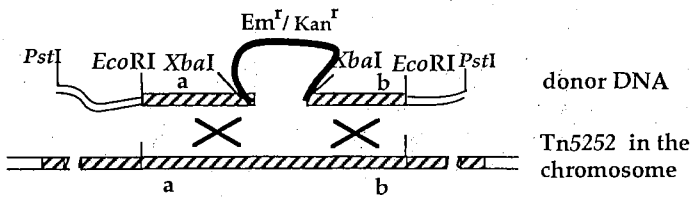
A - Clone the 3.27 kb *Eco*RI fragement of the transposon into pLS1



B - Delete the *Xba*I fragment and Insert a drug resistance marker at that site .



C - Linearize the vector at the *Pst*I, and use it to transform SP1000 (Tn*5252*)



D - The resulting insertion/deletion mutant (Tn*5252* ::Em^r/Kan^r)

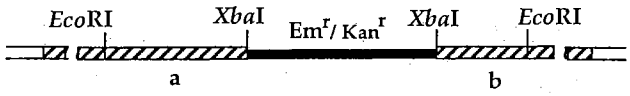


Figure 32. Deletion mutagenesis strategy used to mutate ORF6. The insertion marker was either Em^r (pVA891) or Kan^r (the cassette from pSJ170).

Table XVIII. Transformation experiment to check if the Em insertion is in the transposon.

DONOR DNA	MARKER	Rx1 TRANSFORMANTS /ml	DP1322 TRANSFORMANTS /ml
	str	135000	19200
DP1617	Em	0	0
	str	88000	20500
SP1262	Em	10	280
	str	250000	27000
SP1000 Δ*Xba*I #1	Em	10	700
	str	4900	7100
SP1000 Δ*Xba*I #2	Em	0	40
	str	80000	6800
SP1000 Δ*Xba*I #3	Em	0	1950
	str	102000	5050
SP1000 Δ*Xba*I #4	Em	10	285
	str	19600	3950
SP1000 Δ*Xba*I #5	Em	0	50
	str	7200	2400
SP1000 Δ*Xba*I #6	Em	0	10

One microgram/milliliter of each DNA was used to transform competent Rx1 and DP1322 cells.

One of these strains was designated SP1213 and stored for future use.

Construction of pSJ126

For all the insertion and deletion mutagenesis experiment in the transposon Tn*5252*, pVA891 expressing the erythromycin resistance has been the only reporter used so far. The major limitation to its use as a reporter has been the non-availability of compatible restriction sites in the plasmid. To overcome this problem and to generate a new antibiotic resistance cassette, pSJ126 was generated. As shown in the Fig. 33, pDL276 (53), a streptococcal-*E. coli* shuttle plasmid carrying a kanamycin-resistance gene was digested with *Bss*HII and *Ava*I. Also, pRL425 (56) another vector plasmid carrying ampicillin, erythromycin and chloramphenicol resistance genes, was digested with *Hin*dIII. The two digests were mixed and treated with the Klenow fragment of *E. coli*. After cleaning the DNA, they were ligated and used to transform *E. coli* JM109 cells. Kan^r- Amp^r transformants were selected and replica-plated to check for the absence of the unselected markers, Cm and Em. One of the transformants was designated pSJ126. In pSJ126, eight identical restriction sites flank the *aph*A (Kan^r). Thus, the *aph*A can be released by digesting with any one these 8 enzymes.

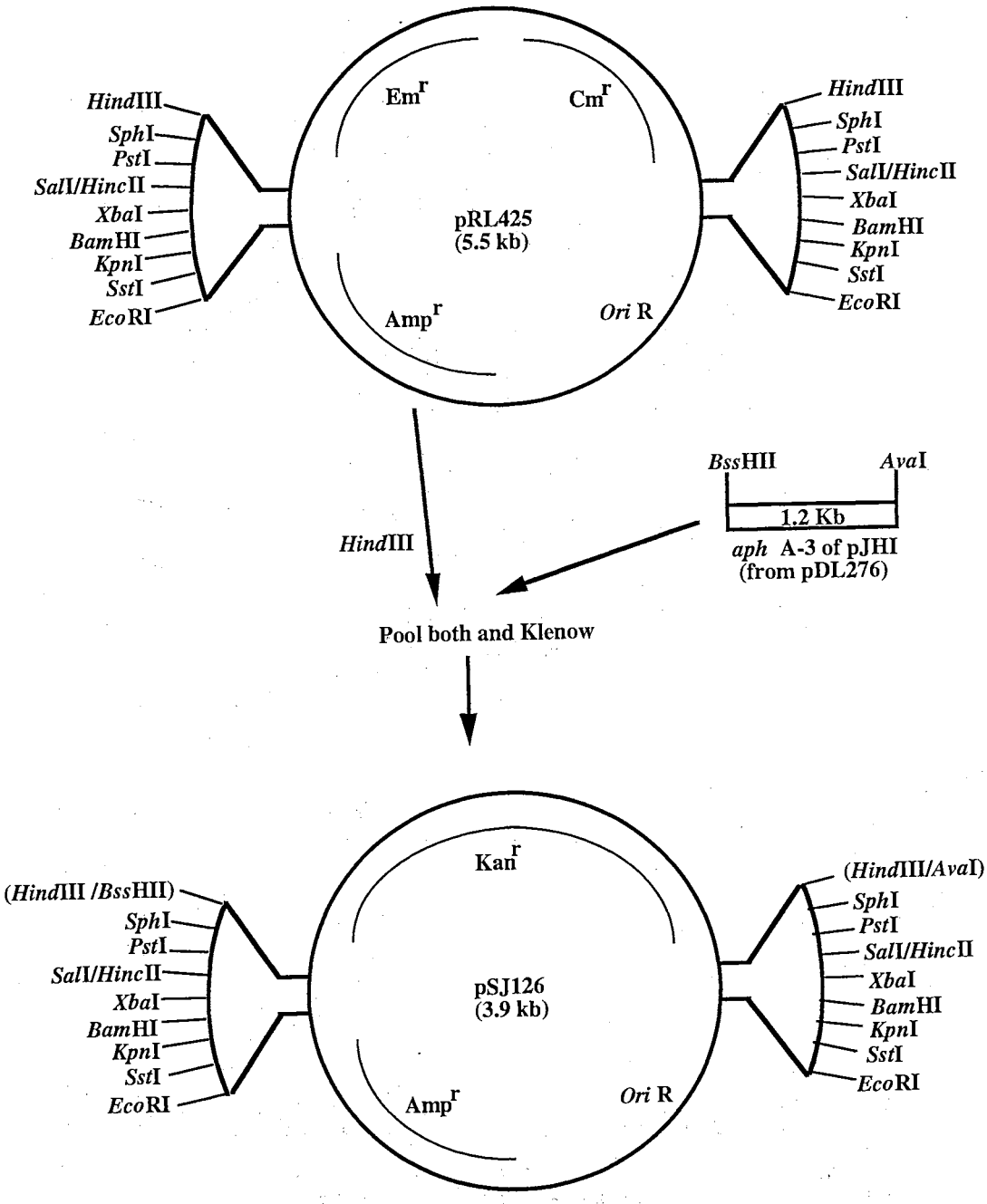


Figure 33. Construction of pSJ126 using pRL425 vector and Kan^r from pDL276.

Construction of pSJ150

Insertional inactivation of the gene of interest in the transposon using the kanamycin cassette from pSJ126 was not satisfactory, as the concentration of kanamycin needed was only slightly higher than the minimum inhibitory concentration to pneumococcus. The reason for this was that the *aph*A obtained from pDL276 was originally from pJH1 (53), a streptococcal plasmid, that was found in *E. fecalis*. Therefore, when introduced as a single copy, the resistance provided to the host was not high. This problem was overcome by using the *aph*A from the conjugative transposon Tn*1545*. It has been shown by others (28) that the *aph*A from pJH1 and Tn*1545* are identical except for some changes in the promoter region. So using the *aph*A from Tn*1545*, another plasmid was constructed. The source of the *aph*A of Tn*1545* was from the pAT21-1 (28). The vector used for this plasmid construction (shown in Fig. 34) was pDG792 (62). pDG792 has the kanamycin gene from pJH1 flanked by 4 unique restriction sites on one side and 16 unique restriction sites on the other side. pDG792 was digested with *Cla*I to release the *aph*A of pJH1 and a 1.6-kb *Cla*I fragment of pAT21-1 containing the *aph*A of Tn*1545* was ligated and used to transform JM109. After transformation, the transformants were selected in LB agar containing 25 μg/ml of kanamycin. When present in multiple copies the *aph*A of Tn*1545* provided the *E. coli* with only a low level of kanamycin resistance whereas, the *aph*A of pJH1 provided the host with high level resistance (50-70 μg/ml of kanamycin). The recombinant plasmid generated was designated pSJ150 and, in this plasmid, 4 unique restriction sites flank

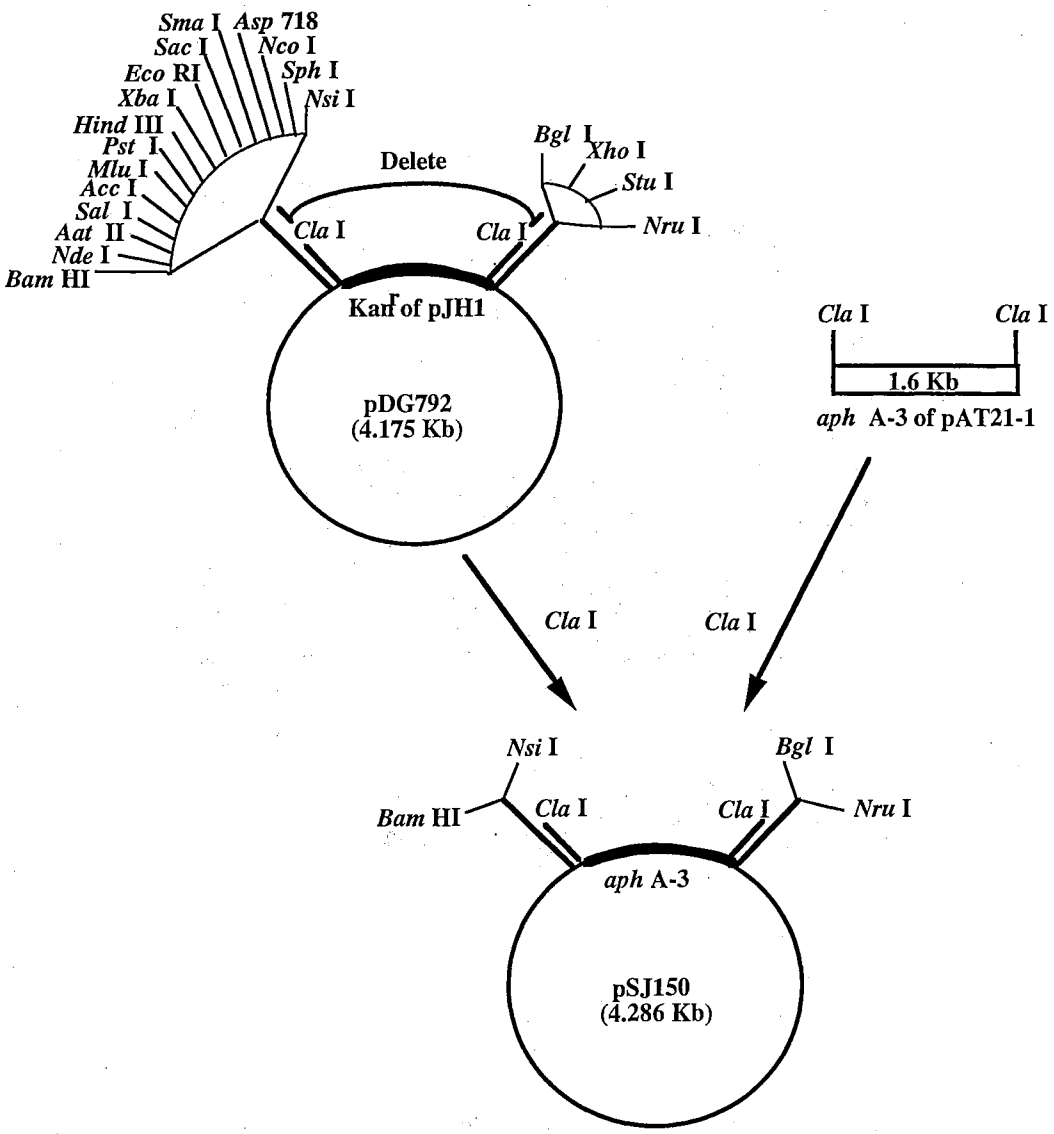


Figure 34. Construction of pSJ150, using pDG792 vector and Kan^r from pAT21-1.

the *aph*Aon one side and 16 unique restriction sites on the other side flank the *aph*A of Tn*1545*. As only one restriction enzyme (*Cla*I) site is present on both sides of the *aph*A, it cannot be released as a cassette by digestion with other enzymes.

Construction of pSJ170

For ease in getting the Tn*1545 aph*A as a cassette for the purpose of insertional inactivation, pSJ170 was constructed using pSJ126 and pSJ150. Construction of pSJ170 is shown in Fig. 35. Briefly pSJ126 was digested with *Hin*cII to release the kanamycin (from pJH1) and a 1.6-kb *Nru*I/*Sma*I fragment of pSJ150 was ligated. *E. coli* JM109 cells were transformed and the recombinant plasmid obtained was designated pSJ170. In pSJ170, ten restriction sites flank the *aph*A on one side and 9 on the other, of this 6 sites are present on both sides that can be used to release the *aph*A as a cassette.

Construction of pSJ172

To create a methylase-deletion-mutant strain carrying Kan^r, pSJ119 was linearized with *Xba*I and ligated to a 1.6 kb *Xba*I fragment of pSJ170. The recombinant plasmid, designated pSJ172 was checked with various restriction enzymes to determine the orientation of the Kan^r insert.

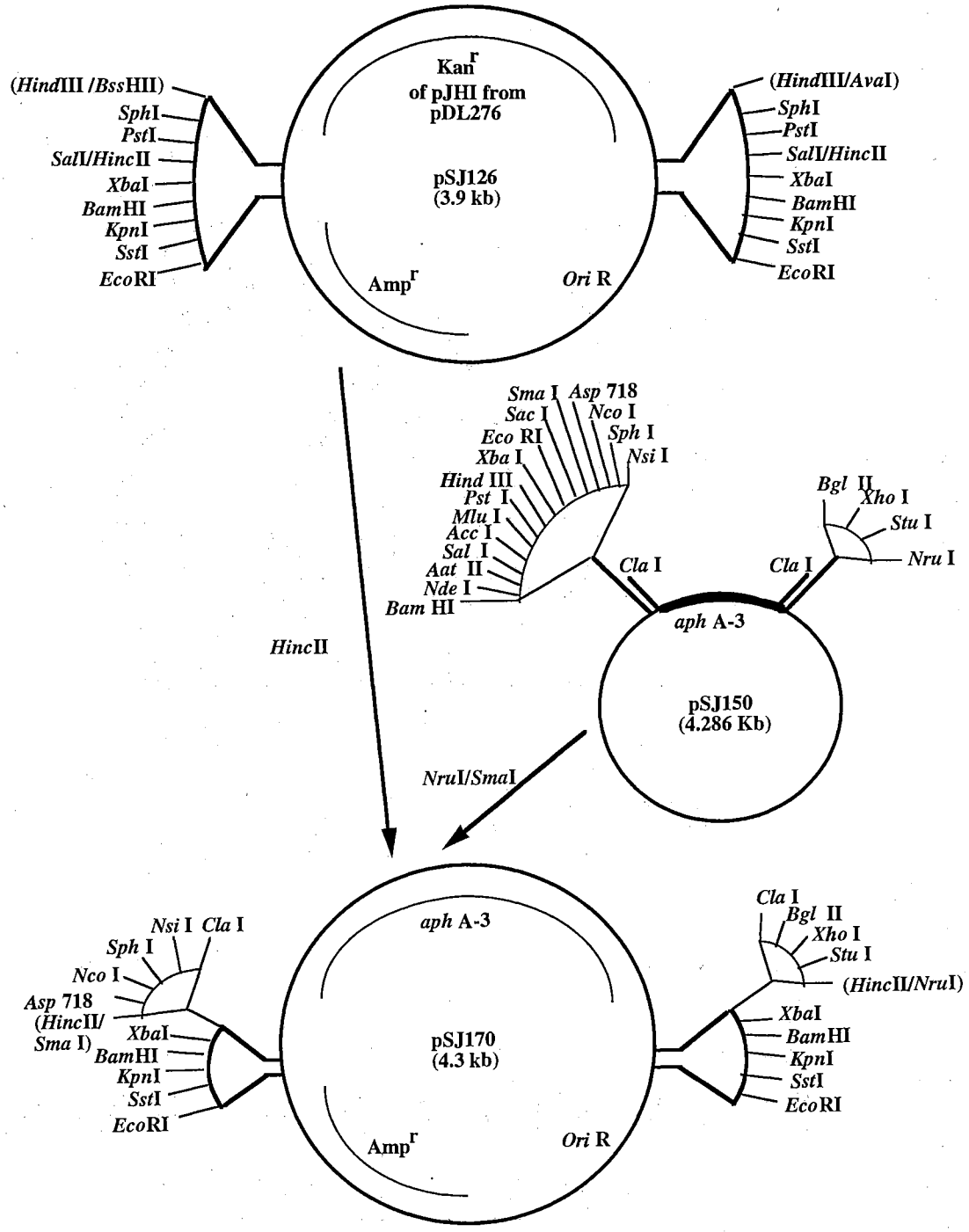


Figure 35. Construction of pSJ170, using pSJ150 and pSJ126.

Construction of *in vivo* deletion mutant (SP1214)

pSJ172 was linearized at the *Pst*I site, present in the pLS1 portion of pSJ119, and used as donor in transformation experiments with SP1000 as recipient. The strategy is as shown in Fig. 32. Kan^r transformants were checked for the unselected markers and one such transformant, designated SP1214 was used for further study. The restriction map of the mutant is shown in Fig. 36. This strain was constructed with the aim to study the transfer of the methylase mutant transposon and the conjugative plasmid, pIP501 (Cm^r, Em^r), from the donor cell to different recipient cells.

Transformation to check if the insertion is in the transposon

Chromosomal DNA from SP1214 was isolated. A asample of the DNA was sheared using a 21 1/2 gauge needle along with the control chromosomal DNAs. Competent Rx1 cells and DP1322 cells were transformed with the sheared and unsheared donor DNA prepared. Transformants were selected on streptomycin (chromosomal marker) and kanamycin (insertion marker) containing agar plates. The results listed in Table XIX indicated that the insertion of kan^r was in the transposon as DP1322 transforms better than Rx1 for kan^r. DP1402, that was used as a control in the transformation experiments carries the kanamycin resistance gene in a conjugative transposon, the BM4200 element, which does not have any homology to Tn*5252*. So no Kan^r transformants were expected to be observed while using this DNA as the donor.

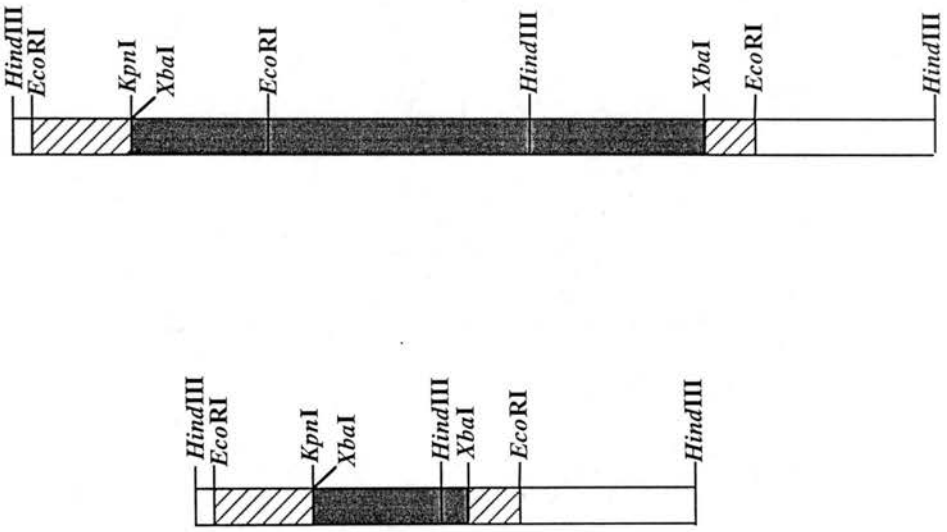


Figure 36. Restriction map of SP1213 and SP1214. In the case of SP1213, pVA891 was inserted after deleting the *Xba*I fragment and in the case of SP1214, Kan^r cassette from pSJ170 was introduced after deleting the *Xba*I fragment. Crosshatched box indicates the region of interest and solid box indicates the inserted marker DNA.

Table XIX. Transformation to show that the insertion in SP1214 is in the transposon.

DONOR DNA	MARKER	Rx1 TRANSFORMANTS /ml	DP1322 TRANSFORMANTS /ml
UNSHEARED	*str*	1250000	250000
DP1617	Kan	0	0
SHEARED	*str*	1300000	230000
	Kan	0	0
UNSHEARED	*str*	38000	30000
SP1214	Kan	0	2940
SHEARED	*str*	48000	13000
	Kan	0	1390
UNSHEARED	*str*	70000	8000
DP1402	Kan	0	0
SHEARED	*str*	31000	5000
	Kan	0	0

One microgram/milliliter of the Donor DNA was used to transform competent Rx1 and DP1322 cells.

Southern Hybridization to check the insertion

Chromosomal DNA from SP1214, SP1213, SP1262, Rx1 and SP1000 were Prepared and digested with *Eco*RI and run a 0.8% agarose gel. The DNA was then transferred to nylon membrane and probed with radiolabled pSJ107. As evident from the autoradiogram, shown in Fig. 37, the probe hybridized to a 2.1 kb fragment in the case of Rx1 (target site) and a 3.27-kb fragment of SP1000 (right junction fragment), as expected. The probe also hybridized to a 5.2 kb and a 3.95 kb fragment in case of SP1262 (pVA891 insertion in the *Kpn*I site). It was expected to hybridize to a 3.1 kb fragment in case of SP1214 as the 1.8 kb *Xba*I fragment is replaced by the 1.6 kb kan^r. In the case of SP1213 the probe was expected to hybridize to a 5.0 kb and 2.4 kb fragment, as pVA891, which has an internal *Eco*RI site, replaces the 1.8 kb *Xba*I fragment. Apart from the presence of these two fragments, an unexpected third band corresponding to a size of 1.1-kb was also seen. The origin of this band is not clear at present.

Creation of SP1206 and SP1207

As mentioned earlier, presence of novobiocin resistance in the recipient during conjugation seems to interfere with the conjugal transfer process. So the chromosomal DNA from the deletion mutants SP1213 and SP1214 (Fig. 36) were used to transform

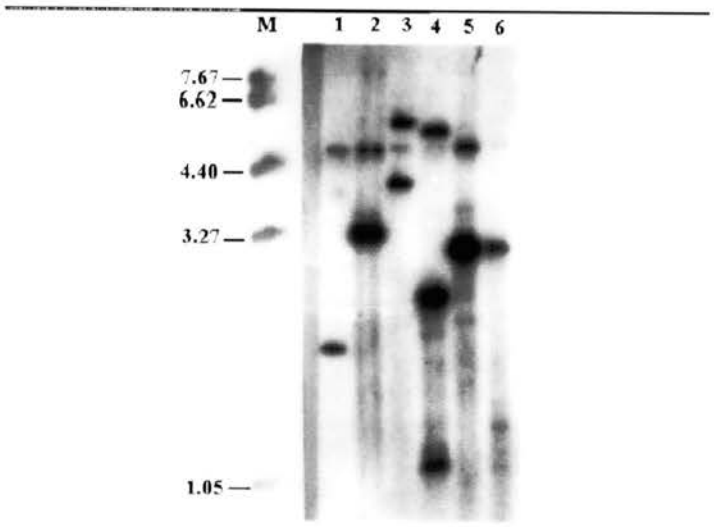


Figure 37. Autoradiogram showing the deletion mutation in SP1213 and SP1214. Chromosomal DNA from Rx1, SP1000, SP1262, SP1213 and SP1214 were isolated, digested with *Eco*RI, run on a 0.8% agarose gel, blotted to a nylon membrane, and probed with radioactively labeled pSJ107. Lane 1, DNA from Rx1; 2, DNA from SP1000 (Rx1::Tn*5252*); 3, DNA from SP1262; 4, DNA from SP1213; 5 and 6 DNA from SP1214 and M, molecular weight markers. Nominal molecular masses are indicated.

competent SP1264 (Tn*5252*, nov^r). The transformants obtained were checked for the presence of appropriate markers. One true transformant designated SP1206 (*nov*^{r}, Cm^r, Em^r, *str*^{s}) and SP1207 (*nov*^{r}, Cm^r, Kan^r, *str*^{s}) from each transformation was stored for future use in filter-mating experiments.

Conjugation of the methylase deletion mutants with *S. pneumoniae* DP1004 (*Dpn*II$^-$) and DP3111 (*Dpn*II$^+$)

The deletion mutants created were used as donors in filter-mating experiments with pneumococcus as the recipient. To check if the transfer was affected by the presence of *Dpn*II restriction-modification system, DP3111 was chosen as one of the recipients. DP1004, which is devoid of any known restriction-modification system, was chosen as the control recipient. The results of the mating, presented in Table XX, shows that the deletion mutation decreases the transfer frequency by one-to-two orders of magnitude. Interestingly, the transfer frequency of SP1207 in the case of DP1004 recipient is reduced by about one order of magnitude, whereas in the case of DP3111, it is reduced by two orders of magnitude. This reduction in the transfer frequency could be due the restriction of the transferring DNA as it is not methylated in the donor prior to transfer.

Table XX. Conjugal transfer frequencies of the deletion mutant to *Dpn*II$^+$ and *Dpn*II$^-$ hosts.

DONOR	RECIPIENT	TRANSCONJUGANTS /DONOR x 10^8
SP1273	DP3111	645 ± 25
SP1206	DP3111	<2
SP1207	DP3111	2 ± 1.4
SP1273	DP1004	119 ± 11
SP1206	DP1004	<2
SP1207	DP1004	4 ± 2

The results are from 2 independent experiments representing the range. SP1273 carries Tn*5252* with pVA891 inserted within the ORF6 and is nov^r. SP1206 carries Tn*5252* with the *Xba*I deletion and the marker DNA inserted at this site is pVA891; SP1207carries Tn*5252* with the *Xba*I deletion and the marker DNA inserted at this site is the kan^r cassette and is nov^r. The recipients used were str^r *S. pneumoniae* that were *Dpn*II$^+$ (DP3111) and *Dpn*II$^-$ (DP1004).

Construction of SP1208

To check the transfer properties of the methylase deletion mutant transposon and pIP501 (a conjugative plasmid from *E. fecalis*) simultaneously, the strain containing both the genetic elements had to be constructed. As the copy number of pIP501 is too low and isolation of the plasmid is difficult, it was conjugally transferred from a donor to the methylase mutant transposon containing cell. SP1207 was used as the recipient and DP3218 was used as the donor. The transconjugants were selected, by the overlay technique, with novobiocin (recipient marker) and erythromycin (plasmid marker). The transfer frequency was expected to be very high (10^{-3} to 10^{-4}); instead the frequency was very low. This low transfer could be due to the presence of *nov* marker in the recipient. However, transconjugants obtained were checked for the unselected markers, Cm^r (transposon, plasmid markers), Kan^r (methylase mutation) and Str^s (donor marker). One true transconjugant (Kan^r, Cm^r, Em^r, Nov^r, Str^s) was designated SP1208 and used for future studies.

Construction of SP1209, SP1210, SP1211, and SP1212

To check for complementation of the methylase deletion mutant, and also to observe any effect due to the presence of pSJ107 or its derivatives on the transfer of the pIP501, SP1208 was transformed with pLS1, pSJ107, pSJ117 and pSJ119. The transformants obtained were checked for the presence of appropriate markers. The strains were designated as SP1209 (SP1208 containing pLS1), SP1210 (SP1208

containing pSJ107), SP1211 (SP1208 containing pSJ117) and SP1212 (SP1208 containing pSJ119) and stored for future use.

Complementation analysis

SP1208 and its derivatives, SP1209, SP1210, SP1211 and SP1212 were used as donors in filter mating experiments with DP3111 as recipient. The transfer of the transposon (methylase-deletion mutant) and the transfer of pIP501 was checked by selecting the transconjugants with streptomycin (recipient marker) and with either kanamycin (for transposon) or with erythromycin (for pIP501). The transconjugants obtained were then checked for the unselected markers as before. As evident from the results shown in Table XXI, in spite of the presence of pSJ107, the transposon was still deficient in transfer. This inability to complement the deletion mutation could be due to the fact that the deletion in the transposon spans not just the amino terminus of ORF6 but also ORF41 and the upstream region which was found to have numerous indirect and direct repeats (refer Fig. 6). These repeats could be involved in regulation of the expression of the entire operon like structure, as –10 and –35 regions were found only for ORF41 and not the others. The Em^r transconjugants obtained from SP1211 indicated that the transfer of pIP501 was complemented to some extent (the normal transfer frequency of pIP501 to a DpnII$^-$ is $10x^{-3}$ to $10x^{-4}$/donor) but the same result was expected from SP1210 donors also as it has pSJ107. The lack of complementation was found to be due

Table XXI. Complementation analysis of the deletion mutants using DP3111 (*Dpn*II$^+$) as the recipient.

DONOR	PLASMID IN DONOR	Em^r TRANSCONJUGANTS /DONORx10^8	Kan^r TRANSCONJUGANTS /DONOR x 10^8
SP1208	pIP501	10 ± 3.2	2 ± 1.4
SP1209	pLS1, pIP501	<5	5 ± 2.2
SP1210	pSJ107,pIP501	<7	<7
SP1211	pSJ117,pIP501	290 ± 17	<7
SP1212	pSJ119,pIP501	190 ± 14	<4

The results are from two independent experiments representing the range. SP1208 is Δ*Xba*I Tn*5252* with Kan^r inserted within the ORF6 and is *nov*^{r}. SP1209 (SP1208 containing pLS1), SP1210 (SP1208 containing pSJ107), SP1211 (SP1208 containing pSJ117) and SP1212 (SP1208 containing pSJ119) were also used as donors. The recipient used was *str*^{r} *S. pneumoniae* that was *Dpn*II$^+$ (DP3111).

to the donor itself. Analysis of SP1209, SP1210 showed that the strains were not Em^r but Em^s, indicting that pIP501 was lost from the donor prior to or during transfer. However, these experiments demonstrate that the methylase deletion mutation of the transposon could not be complemented.

CHAPTER V

DISCUSSION

Studies directed towards understanding the sudden worldwide emergence of multiple-antibiotic-resistant *S. pneumoniae* led to the identification of a conjugative transposon, Tn5253 (176). Preliminary studies showed that Tn*5253* was a composite transposon consisting of Tn*5251* and Tn*5252* (7). Tn*5252* is the prototype of the larger elements and our lab has focused on understanding the biology of Tn*5252*. The ends of this transposon have been shown to be essential for conjugal transfer (97). Work on the left end and the middle region of the element have been reported earlier (1, 2, 188, 189). The current work was undertaken to analyze the right end of the transposon at the molecular level. To achieve this, attempts were made to clone the 3.27 kb right-junction fragment in its entirety. The fragment could not be cloned in *E. coli* but was clonable in *S. pneumoniae*. This indicated that some protein or proteins specified by the transposon DNA were toxic to the *E. coli* host. The nucleotide sequence of the fragment of interest was obtained using subclones and the sequences were manually assembled. Analysis of the sequence revealed the presence of four open-reading frames (ORF41, ORF6α, ORF6β, and ORF42), all transcribed in the same orientation. ORF6β was localized within ORF6α but in an alternate reading frame.

In vitro transcription-translation studies revealed the presence of at least two of the four expected proteins. The sizes of these two proteins matched those of ORF6α and ORF6β. Genbank search showed that the amino-acid sequences of ORF6α were similar to a variety of 5^{m}C-cytosine methyltransferases. DNA methyltransferases are a universally distributed, diverse group of enzymes and are a part of the restriction-modification systems. The restriction-modification systems have been divided into three types, TypeI, Type II and Type III. Some of the properties of these three types are listed in Table XXII. Methyltransferases belonging to these three types have been shown to play essential roles in a number of biological processes including gene regulation, DNA mismatch repair, and restriction-modification (131, 143). Type II methylases have three major groups and they are 5^{m}C-cytosine methylases, 4^{m}N-cytosine methylases, and 6^{m}N-adenine methylases. Among the three major groups, there is a remarkable amino acid similarity between the 5^{m}C-cytosine methylases and a remarkable structural similarity between all of the type II methylases (129). The deduced amino acid sequence of ORF6α contained all the ten conserved motifs found in prokaryotic Type II methylases and the order of the motifs was the same as those of 5^{m}C-cytosine methylases. With the data from *in vitro* methylation experiments, we have demonstrated the methylating activity of protein extracts of the cells carrying the ORF6α gene. However, the highly purified ORF6α fusion protein obtained from the insoluble fraction of these cells lacked methylase activity implying that the denaturation and renaturation steps involved in the isolation may have led to the loss of the enzymatic activity. On the other hand, it is possible that there was a need for as yet unidentified additional factors for methylation

Table XXII. Some Properties of the three Types of Restriction-Modification systems.

TYPE	STRUCTURE	CO-FACTOR REQUIREMENTS R	M	SITE OF METHYLATION/ CLEAVAGE
I	Enzyme complex consisting ENase, MTase and specificity determining subunits	ATP, SAM, Mg^{++}	SAM (ATP, Mg^{++})	Remote from the Target site
II	ENase and MTase are separate enzymes	Mg^{++}	SAM	Within the target or at a defined location within 20 bp on one site of the target
III	Enzyme complex consisting of ENase and MTase	ATP, SAM, Mg^{++}	SAM (ATP, Mg^{++})	At a defined distance of 25-27 bp from the 3' end of the target

ENase refers to the endonuclease and MTase to Methylase. Cofactors shown in parentheses indicate that their presence stimulates the methylation (M) reaction. The above information was obtained from reference 143.

activity by the ORF6α product. It is worth noting nere the unusual structure of the DNA cytosine methylases, M.*Eco*HK311 (122) and M.*Aqu*I (95), which require the expression of two polypeptides for enzymatic activity. Interestingly in case of M.*Eco*HK311, one of the polypeptides, the β sub-unit, is encoded within an alternate reading frame of the polypeptide α (122). ORF6β is also encoded within an alternate reading frame of ORF6α and it showed a significant sequence similarity of the β sub-unit of M.*Eco*HK311. However in the case of M.*Eco*HK311, the α or the β sub-unit alone has been shown to be functional. Addition of the α to the β sub-unit has been shown to increase the methylation potential of the enzyme, M. *Eco*HK311, by about ten fold. In the case of M.*Eco*HK311, conserved motif IX, absent in the α-subunit was found to be present in the β sub-unit. In the case of ORF6β, no conserved sequence motif normally found in the methylase was present. Also, the region of homology between ORF6β and the β sub-unit of M.*Eco*HK311 did not span the conserved motif. Unlike the α sub-unit of M.*Eco*HK311, the insoluble- fraction-purified ORF6α was not functional. Attempts to clone ORF6β in expression vectors to over-express and purify the protein were unsuccessful. So attempts were made to purify the functional methylase from the soluble fraction, as the extract of the soluble fraction was found to be active. The partially-purified methylase fraction from the soluble fraction using affinity chromatography yielded a mixture of several proteins of various molecular masses. Nevertheless, this fraction showed an increase in the specific activity of the methylase. Attempts to further purify the methylase from this mixture using HPLC involving size-exclusion chromatography were unsuccessful. Analyses of the HPLC data revealed that the

proteins were forming aggregates and thus eluting in the first few fractions. However, the fractions collected failed to show any methylase activity. The reason for the disappearance of methylase activity is not known at present. Having not resolved this, the possibility of two or more proteins interacting with one another to form a functional methylase is still an open question.

Comparison of the structure of the three classes of type II methylases (the 4^{m}N-cytosine methylase, 6^{m}N-adenine methylase and the 5^{m}C-cytosine methylase) has revealed the presence of nine of the ten conserved motifs in the ORF. The conserved motif corresponding to motif IX of 5^{m}C-cytosine methylases is absent in 4^{m}N-cytosine methylases and in 6^{m}N-adenine methylases. The α sub-unit of M.*Eco*HK311 (a 5^{m}C-cytosine methylase) has been shown to function in spite of lacking motif IX suggesting that ORF6α protein purified from the insoluble fraction should be functional, at least to some extent, in the absence of ORF6β. In addition, the presence of a protein of the size expected for ORF6β was observed in the insoluble fraction purified methylase. The presence of all ten conserved motifs in ORF6α, the lack of any conserved element in ORF6β and the lack of any methylase activity in the insoluble fraction purified methylase strongly suggests that ORF6β might not be involved in the process. Thus, the loss of activity could be due to the denaturation-renaturation process itself.

In any event, this is the first DNA 5^{m}C-cytosine methylase known in pneumococci and also the first methylase to be identified in a transposon. The Tn*5252* methylase has been named M.*Spn5252*I. Analysis of the methylase activity of the pneumococcal host carrying the transposon indicated that the activity is similar to that of a wild-type pneumococcal strain. Also, there are several inverted and direct repeats

observed upstream of ORF6 suggesting that the methylase might be regulated. Attempts to identify the time of induction of the methylase during mating using antibodies directed against ORF6α has so far been unsuccessful. This could be due to the fact that the induction of the methylase might be too low to detect as only very few molecules are needed for the methylation process. The presence of a large variable region between the conserved motif VIII and IX indicated that the methylase might recognize novel target sequences for methylation suggesting that M.*Spn5252*I might be a member of the multi-specific family of methylases. Based on the *in vitro* methylation assays protection against *Eco*RII restriction, *Hae*III restriction and *Sau*3AI restriction was observed. Protection of *Sau*3AI site is of evolutionary importance as this is the site the *Dpn*II (for *Diplococccus pneumoniae*) restriction system recognizes for restriction. Thus, as in the case of bacteriophages of *B. subtilis*, Tn*5252* has a protective mechanism to overcome the host's potential to restrict unmethylated GATC site. M.*Spn5252*I has also been shown to protect the DNA against restriction by *Eco*RII and *Hae*III. Restriction systems directed towards unmethylated *Eco*RII or *Hae*III sites in streptococci are not known at present but the presence of a protective mechanism against these two sites in the streptococcal transposon, Tn*5252* suggests that this restriction could possiblly be there.

Guild, *et al.* (64) reported that Tn*5253*, the parental element of Tn*5252* (7), was not subjected to restriction during transfer by conjugation to a *Dpn*II$^+$ strain while the simultaneous transfer of conjugative plasmid, pIP501, was reduced 10^4 fold. This implies that Tn*5253* was in a protected state only during the mating event came from the observation that the chromosomal DNA was sensitive to *Dpn*II digestion *in vitro*. M.*Spn5252*I might be the one responsible for this protection as it has been shown to give

protection against *Sau*3AI, which is an isoschizomer of *Dpn*II. The ten-fold difference in the transfer frequency of the *Xba*I-deletion mutant, between a *Dpn*II$^-$ and a *Dpn*II$^+$ recpient also confirms the protective role of the methylase. To study the role of the methylase, an insertion mutant was created which showed hyperactivity. This increased activity might be due to the removal of the specificity domain because of the frame shift mutation introduced at the carboxyl end. The deletion mutant constructed here reduced the transfer frequency dramatically but the deletion spans an upstream ORF and the regulatory regions also. Complementation analysis revealed that this deletion mutant could not be complemented, by the presence of the 3.27 kb fragment in a plasmid in trans. This further supports the idea that the deficiency in transfer is due to loss of the regulatory signals. Thus, in this study I have obtained the nucleotide sequence of the 3.5 kb right junction fragment, demonstrated the presence of a functional 5^mC cytosine methyltransferase, identified some of the target sites for methylation, and showed that the methylase is involved in anti-restriction activity.

Analysis of the methylase to determine if it is made up of two sub-units and the role of the methylase in regulation awaits further study. One way to study this would be to purify the methylase from the insoluble fraction as before but instead of eluting the protein and then doing the step dialysis to remove the urea, it can be done in the column itself. This procedure would not give adequate time for the proteins to come together and form aggregates like dialysis. If the protein eluted after removing urea in the column shows activity then one can conclude that ORF6β or any other protein is not needed for the methylase activity, as urea would prevent protein interactions. Another interesting question that needs to be answered concerns the regulation of M.*Spn5252*I during

conjugal transfer. Western blots could be done using the lysates obtained by mating a donor carrying pSJ107 and Tn*5252* with a wild-type pneumococcal recipient. This would ease visualization of the methylase as more methylase would be made from the plasmid. Regulation of the methylase could also be studied by using *lacZ* reporter constructs. Apart from playing a protective role, the methylase might also be responsible for regulating the transfer process of the transposon as in the case of Tn*5* (222) and IS*10* (157). This question could be answered only by constructing strains that carry deletion mutations totally within ORF6 and not the flanking regions.

The definitively demonstrated roles of DNA methylases in X-chromosome inactivation, oncogenesis, DNA mismatch repair, and protection from restriction show the biological importance of this class of enzymes. Further studies should enlarge our understanding of the biological roles of these interesting molecules.

REFRENCES

1. **Alarcon-Chaidez, F. J.** 1997. Genetic and transcriptional organization of a region essential for DNA transport in Tn*5252*. Ph. D. Thesis. Department of Microbiology and Molecular Genetics. Oklahoma State University.

2. **Alarcon-Chaidez, F. J., J. Sampath, P. Srinivas, and M. N. Vijayakumar.** 1996. Tn*5252*: A model for complex Streptococcal Conjugative transposons. p. 1029-1032. In "*Streptococci* and the host". Horaud, T., A. Bouvet, R. Leclercq, H. DeMontclos, and M. Sicard (eds.). Plenum Publishing, New York.

3. **AlonsoDeVelasco, E., A. F. M. Verheul, J. Verhoef, and H. Snippe.** 1995. Streptococcus pneumoniae: Virulence Fcators, Pathogenesis, and vaccines. Micrbiol. Rev. **59:** 591-603.

4. **Amann, E., and J. Borsius.** 1985. 'ATG' vectors for regulated high-level expression of cloned genes in *Escherichia coli*. Gene. **40:** 183-190.

5. **Avery, O. T., C. M. Macleod, M. McCarty.** 1944. Studies on the chemical nature of the substance inducing transformation of pneumococcal serotypes: induction of transformation by a desoxyribonuceleic acid fraction isolated from pneumococcus type III. J.Exp.Med. **79:** 137-157.

6. **Ayalew, S.** 1994. Molecular Basis for the Target Specificity and Insertion of the Conjugative Transposon Tn*5252*. Ph. D. Thesis. Department of Microbiology and Molecular Genetics. Oklahoma State University.

7. **Ayoubi, P., A. O. Kilic, and M. N. Vijayakumar.** 1991. Tn*5253*, the pneumococcal (*cat tet*) BM6001 element, is a composite structure of two conjugative transposons, Tn*5251* and Tn*5252*. J. Bacteriol. **173:** 1617-1622.

8. **Bachi, B., J. Reiser, and V. Pirrotta.** 1979. Methylation and cleavage sequences of the EcoPi restriction-modification enzyme. J. Mol. Biol. **128:** 143-163.

9. **Balganesh, T. S., L. Reiners, R. Lauster, M. Noyer-Weidner, K. Wilke, and T. A. Trautner.** 1987. Construction and use of chimeric SPR/ϕ3T DNA methyltransferases in the definition of sequence recognizing enzyme regions. EMBO J. **6:** 3543-3549.

10. **Bandyopadhyay, P. K., F. W. Studier, D. L. Hamilton, and R. Yuan.** 1985. Inhibition of the type I restriction-modification enzymes *Eco*B and *Eco*K by the gene 0.3 protein of bacteriophage T7. J. Mol. Biol. **182:** 567-578.

11. **Barany, F., M. Danzitz, J. Zebala, and A. Mayer.** 1992. Cloning and sequencing of genes encoding the *Tth*HB8I restriction and modification enzyme: comparison with isoschizomeric *Taq*I enzymes. Gene. **112:** 3-12.

12. **Belogurov, A. A., E. P. Delver, O. V. Rodzevich.** 1992. IncN plasmid pKM101 and IncI1 plasmid ColIb-P9 ecode homologous antirestriction proteins in their leading regions. J. Bacteriol. **174:** 5079-5085.

13. **Belogurov, A. A., T. N. Yussifiv, V. U. Kotova, and G. B. Zavigelsky.** 1985. The novel gene(s) ARD of plasmid pKM101: alleviation of EcoK restriction. Mol. Gen. Genet. **198:** 509-513.

14. **Berheimer, H. P.** 1977. Lysogeny in pneumococci freashly isolated from man. Science. **195:** 66-67.

15. **Berkner, K. L., and W. R. Folk.** 1977. *Eco*RI cleavage and methylation of DNAs containing modified pyrimidines in the recognition sequence. J. Biol. Chem. **252:** 3185-3193.

16. **Berkner, K. L., and W. R. Folk.** 1979. The effects of substituted pyrimidines in DNAs on cleavage by sequence-specific endonucleases. J. Biol. Chem. **254:** 2551-2560.

17. **Bickle, T. A., and D. H. Kruger.** 1993. Biology of DNA restriction. Microbiol. Rev. **57:** 434-450.

18. **Biggin, M. D., T. G. Gibson, and G. F. Hong.** 1983. Buffer gradient gels and ^{35} S label as an aid to rapid DNA sequence determination. Proc. Natl. Acad. Sci. USA. **80:** 3963-3965.

19. **Bradford, M. M.** 1976. A Rapid and Sensitive Method for the Quantitation of MicroGram Quantities of Protein Utilizing the Principle of Protein-Dye Binding.Anal. Biochem. **72:** 248-254.

20. **Brockes, J. P., P. R. Brown, and K. Murray.** 1972. The deoxyribonucelic acid modification enzyme of bacteriophage P1. Biochem. J. **127:** 1-10.

21. **Bron, S., E. Luxen, and G. Venema.** 1983. Resistance of bacteriophage H1 to restriction and modification by Bacillus subtilis R. J. Virol. **46:** 703-708.

22. **Bron, S., L. Janniere, S. D. Ehrlich.** 1988. Restriction and modification in *Bacillus subtilis* Marburg 168: target sites and effects on plasmid transformation. Mol. Gen. Genet. **211:** 186-189.

23. **Buhk, H. J., B. Behrens, R. Tailor, K. Wilke, J. J. Prada, U. Gunthert, M. Noyer-Weidner, S. Jentsch, and T. A. Trautner,** 1984. Restriction and modification in *Bacillus subtilis*: nucleotide sequence, functional organization and product of the DNA methyltransferase gene of bacteriophage SPR. Gene **29:** 51-61.

24. **Burckhardt, J., J. Weisemann, and R. Yuan.** 1981. Characterization of the DNA methylase activity of the restriction enzyme from *Escherichai coli* K. J. Biol. Chem. **256:** 4024-4032.

25. **Burdett, V.** 1980. Identification of tetracycline-resistant R-plasmids in *Streptococcus agalactiae* (group B). Antimicrob. Agents Chemother. **18:** 753-760.

26. **Buu-Hoi, A., and T. Horodniceanu.** 1980. Conjugative transfer of multiple antibiotic resistance markers in *Streptococcus pneumoniae* J. Bacteriol. **143:** 313-320.

27. **Caillaud, F., and P. Courvalin.** 1987. Nucleotide sequence of the ends of the conjugative shuttle transposon Tn*1545*. Mol.Gen. Genet. **209:**110-115.

28. **Caillaud, F., P. Trieu-Cuot, C. Carlier, and P. Courvalin.** 1987. Nucleotide sequence of the kanamycin resistance determinant of the pneumococcal transposon Tn*1545*: evolutionary relationships and transcriptional analysis of *aph*A-3 genes. Mol Gen Genet. **207:** 509-513.

29. **Carlier, C., and P. Courvalin.** 1982. Resistance of streptococci to aminoglycoside-aminocyclitol anitbiotics. pp. 162-166. In "Microbiology-1982", D. Schlessinger (ed.), American Society for Microbiology, Washington, D. C.

30. **Chandrasegaran, S., and H. O. Smith.** 1987. Amino acid sequence homologies among twenty-five restriction endonucleases and methylases, p. 149-156. In "Structure and Expression. Froms proteins to ribsomes", R. H. Sarma and M. H. Sarma (ed.), Adenine Press, Guilderland, New York.

31. **Chang, A. C. Y., and S. Cohen.** 1978. Construction and characterization of amplifiable multicopy DNA cloning vehicles derived from the P15A cryptic miniplasmid. J. Bacteriol. **134:** 1141-1156.

32. **Chen, L., A. M. MacMillan, W. Chang, K. Ezaz-Nikpay, W. S. Lane, and G. L. Verdine.** 1991. Direct identification of the active-site nucelophile in a DNA (cytosine-5)-methyltransferase. Biochemistry. **30:** 11018-11025.

33. **Cheng, X.** 1995. Structure and function of DNA methyltransferases. Annu. Rev. Biophys. Biomol. Struct. **24:** 293-318.

34. **Cheng, X., S. Kumar, J. Posfai, J. W. Pflugrath, and R. J. Roberts.** 1993. Crystal structure of the *Hha*I methylasecomplexed with S-adenosyl-L-methionine. Cell. **74:** 299-307.

35. **Clewell, D. B.** 1981. Plasmids, Drug resistance, and gene transfer in the genus *Streptococcus*. Microbiol. Rev. **45:** 409-436.

36. **Clewell, D. B.** 1985. Sex pheromones, plasmids, and conjugation in *Streptococcus faecalis*. pp. 13-28.In " The origin and evolution of Sex", H. O. Halvorson, A. Monroy (eds.). Alan R. Liss, New York.

37. **Clewell, D. B., and C. Gawron-Burke.** 1986. Conjugative transposons and the dissemination of antibiotic resistance in streptococci. Ann. Rev. Microbiol. **40:** 635-659.

38. **Clewell, D. B., and S. E. Flannagan.** 1993. The conjugative transposons of Gram positive bacteria. pp. 369-393. In "Bacterial conjugation", D. B. Clewell (ed.), Plenum Press, New York.

39. **Clewell, D. B., F. An, B. A. White, and C. Gawron-Burke.** 1985. *Streptoccocus faecalis* sex pheromone (cAM373) also produced by *Staphylococcus aureus* and identification of a conjugative transposon (Tn*918*). J. Bacteriol. **162:** 1212-1220.

40. **Clewell, D. B., P. K. Tomich, M. C. Gawron-Burke, A. E. Franke, Y. Yagi and F. Y. An.** 1982. Mapping of Streptococcus faecalis plasmids pAD1 and pAD2 and studies relating to transposition of Tn*917*. J. Bacteriol. **152:** 1220-1230.

41. **Clewell, D. B., S. E. Flannagan and D. D. Jaworski.** 1995. Unconstrained bacterial promiscuity: the Tn*916*-Tn*1545* family of conjugative transposons. Trends in Microbiology. **3:** 229-236.

42. **Clewell, D. B., S. E. Flanagan, Y. Ike, J. J, Jones, C. Gawron-Burke.** 1988. Sequence analysis of termini of conjugative transposon Tn*916*. J. Bacteriol. **170:** 3046-3052.

43. **Connaughton, J. F., W. D. Kaloss, P. G. Vanek, G. A. Nardone, and J. G. Chirikjian.** 1990. The complete sequence of the *Bacillus amyloliquifaciens* proviral H2, *Bam*HI methylase gene. Nucleic Acids Res. **18:** 4002.

44. **Courvalin, P., and C. Carlier.** 1987. Tn*1545*: a conjugative shuttle transposon. Mol. Gen. Genet. **206:** 259-264.

45. **Courvalin, P., C. Carlier, and F. Caillaud.** 1987. Functional anatomy of the conjugative shuttle transposon Tn*1545*. pp. 61-64. In "Streptococcal genetics", J. J. Ferretti and R. Curtiss III. (ed.), American Society for Microbiology. Washington, D. C.

46. **Dagan, R., M. Isaachson, R. Lang, J. Karpuch, C. Block, and J. Amir.** (for the Israeli Pediatric Bacteremia and Meningitis Group). 1994. Epidemiology of pediatric meningitis caused by *Haemophilus influenzae* type b, *Streptococcus pneumoniae*, and *Neisseria meningitidis* in Israel: a 3-year nationwide prospective study. J. Infect. Dis. **169:** 912-916.

47. **Dang-Van, A., G. Tiraby, J. F. Acar, W. V. Shaw, and D. H. Bouanchaud.** 1978. Chloramphenicol resistance in *Streptococcus pneumoniae:* enzymatic acylation and possible plasmid linkage. Antimicrob. Agents Chemother. **13:** 577-582.

48. **Davis, M. A., R. W. Simons, and N. Kleckner.** 1985. Tn*10* protects itself at two levels from fortuitous activation by external promoters. Cell. **43:**379-387.

49. **Delver, E. P., V. U. Kotova, G. B. Zavigelsky, and A. A. Belogurov.** 1991. Nucleotide sequence of the gene (*ard*) encoding the anitrestriction protein of plasmid ColIb-P9. J. Bacteriol. **173:** 5887-5892.

50. **Dixon, J. M. S.** 1967. Pneumococcus resistant to erythromycin and lincomycin. Lancet. **1:** 573.

51. **Dodson, K. W., and D. E. Berg.** 1989. Factors affecting transposition activity of IS*50* and Tn*5* ends. Gene. **76:** 207-213.

52. **Dunn, J. J., and F. W. Studier.** 1983. The complete nucelotide sequence of bacteriophage T7 DNA, and the locations of T7 genetic elements. J. Mol. Biol. **166:** 477-535.

53. **Dunny, G. M., L. N. Lee, and D. J. LeBlanc.** 1991. Improved electroporation and cloning vector system for Gram-positive bacteria. Appl. Envtl. Microbiol. **57:** 1194-1201.

54. **Duplay, P., H. Bedouelle, A. Fowler, I. Zabin, W. Saurin, and M. Hofnung.** 1984. Sequences of the *mal*E gene and of its product, the maltose-binding protein of *Escherichia coli* K12. J. Biol. Chem. **259:** 10606-10613.

55. **Efimova, E. P., E. P. Delver, and A. A. Belogurov.** 1988. Alleviation of type I restriction in adenine methylase (*dam*) mutants of *Escherichia coli*. Mol. Gen. Genet. **214:** 313-316.

56. **Elhai, J., and P. Wolk.** 1988. A versatile class of positive-selection vectors based on the nonviability of palindrome containing plasmids that allows cloning into long polylinkers. **68:** 119-138.

57. **Evans, W., and D. Hansman** 1963. Tetracycline-resistant pneumococcus. Lancet. **1:** 451.

58. **Foster, T., M. A. Davis, K. Takeshita, D. E. Roberts, and N. Kleckner.** 1981. Genetic organization of transposon Tn*10*. Cell. **23:** 201-213.

59. **Franke, A. E., and D. B. Clewell.** 1981. Evidence for a chromosome-borne resistance transposon (Tn*916*) in *Streptococcus faecalis* that is capable of conjugal transfer in the absence of a conjugative plasmid. J. Bacteriol. **145:** 494-502.

60. **Friedman, S., and N. Ansari.** 1992. Binding of the *Eco*RII methyltransferase to 5-fluorocytosine-containing DNA. Isolation of a bound protein. Nucleic Acids Res. **20:** 3241-3248.

61. **Graves, M. C., and C. Rabinowitz.**1986. *In vivo* and *in vitro* transcription of the *Clostridium pasteurianum* ferridoxin gene. J. Biol. Cehm. **261:** 11409-11415.

62. **Guerout-Fleury A.M., K. Shazand, N. Frandsen, P. Stragier.** 1995. Antibiotic-resistance cassettes for *Bacillus subtilis*. Gene. **167:** 335-336.

63. **Guild, W. R., and N. B. Shoemaker.** 1976. Mismatch correction in pneumococcal transformation: donor length and *hex*-dependent marker efficiency. J. Bacteriol. **125:**125-135.

64. **Guild, W. R., M. D. Smith, and N. B. Shoemaker.** 1982. Conjugative transfer of chromosomal R determinants in *Streptococcus pneumoniae*, p. 88-92. In "Microbiology-1982", D. Schlessinger (ed.). American Society for Microbiology, Washington, D.C.

65. **Guild, W. R., S. Hazum, and M. D. Smith.** 1981. Chromosomal location of conjugative R determinants in strain BM4200 of *Streptococcus pneumoniae*. p. 610. In "Molecular biology, pathogenecity, and ecology of bacterial plasmids", S. B. Levy, R. C. Clowes, and E. L. Koenig (ed.). Plenum publishing Corp., New York.

66. **Günthert, U., and L. Reiners.** 1987. *Bacillus subtilis* phage SPR codes for a DNA methyltransferase with triple sequence specificity. Nulceic Acids Res. **15:** 3689-3702.

67. **Günthert, U., R. Lauster, and L. Reiners.** 1986. Multispecific DNA methyltransferases from *Bacillus subtilis* phages. Properties of wild-type and various mutant enzymes with altered DNA affinity. Eur. J. Biochem. **159:** 485-492.

68. **Hadi, S. M., B. Bachi, J. C. W. Shepard, R. Yuan, K. Ineichen, and T. A. Bickle.** 1979. DNA recognition and cleavage by the *Eco*P15 restriction endonuclease. J. Mol. Biol. **134:** 655-666.

69. **Hadi, S. M., B. Bachi, S. Iida, and T. A. Bickle.** 1983. DNA restriction-modification enzymes of phage P1 and plasmid P15B: subunit functions and structural homologies. J. Mol. Biol. **165:** 19-34.

70. **Hamburger, M, Jr., L. H. Schmidt, C. L. Sesler, J. M. Ruegsegger, and E. S. Grupen.** 1943. The occurrence of Sulfonamide-resistant pneumococci in clinical practice. J. Infect. Dis. **73:** 12-30.

71. **Hanahan, D.** 1983. Studies on transformation of *Escherichia coli* with plasmids. J. Mol. Biol. **166:** 557-580.

72. **Hansman, D., and M. M. Bullen.** 1967. A resistant pneumococcus. Lancet. **2:** 264.

73. **Hartley, D. L., K. R. Jones, J. A. Tobian, D. J. LeBlanc, and F. L. Macrina.** 1984. Disseminated tetracycline resistance in oral streptococci: implication of a conjugative transposon. Infect. Immun. **45:** 13-17.

74. **Hatmann, S.** 1982. DNA methyltransferase-dependant transcription of the phage Mu *mom* gene. Proc. Natl. Acad. Sci. USA. **79:** 5518-5521.

75. **Hemphill, H. E., and Whiteley, H.R.** 1975. Bacteriophages of *Bacillus subtilis*. Bacterio. Rev. **39:** 257-315.

76. **Highlander, S. K., and O. Garza.** 1996. The restriction-modification of *Pasteurella haemolytica* is a member of a new family of type I enzyme. Gene. **178:** 89-96.

77. **Hill, C., L. A. Miller, T. R. Klaenhammer.** 1991. *In vivo* genetic exchange of a functional domain from a type II a methylase between lactococcal plasmid pTR2030 and a virulent bacteriophage. J. Bacteriol. **173:** 4363-4370.

78. **Hochhut, B., K. Jahreis, J. W. Lengeler, K. Schmid.** 1997. CTn*scr94*, a conjugative transposon found in Enterobacteria. J. Bacteriol. **179:** 2097-2102.

79. **Horn, N., H. M. Dodd, and M. J. Gasson.** 1991. Nisin biosynthesis genes are encoded by a novel conjugative transposon. Mol. Gen. Genet. **228:** 129-135.

80. **Horodniceanu, T., L. Bougueleret, G. Beith.** 1981. Conjugative transfer of multiple-antibiotic resistance markers in β hemolytic group A, B, D, and G streptococci in the absence of extrachromosomal deoxyribonucleic acid. Plasmid. **5:** 127-137.

81. **Hotchkiss, R. D.** 1951. Tranfer of penicliilin resistance in pneumococci by the dezoxyribonucleate derived from resistant cultures. Cold. Spr. Harb. Symp. Quant. Biol. **16:** 457-461.

82. **Hotchkiss, R. D., and J. Marmur.** 1954. Double marker transformation as evidence of linked factors in Desoxyribonulceate transforming agents. Proc. Natl. Acad. Sci. USA. **40:** 55-60.

83. **Inamine, J., and V. Burdett.** 1982. The genetic and physical characterization of two chromosomal resistance determinants of *Streptococcus agalactiae*. pp. 170-

173. In "Microbiology-1982", D. Schlessinger (ed.). American Society for Microbiology, Washington, D.C.

84. **Inamine, J., and V. Burdett.** 1985. Structural organization of a 67 kilobase streptococcal conjugative element mediating muliple antibiotic resistance. J. Bacteriol. **161:** 620-626.

85. **Ingrosso, D., A. V. Fowler, J. Bleibaum, and S. Clarke.** 1989. Sequence of the D-aspartyl/ L-isoaspartyl protein methyltransferase from human erythrocytes, common sequence motifs for protein, DNA, RNA, and small molecular S-adenosylmethinine-dependent methyltransferases. J. Biol. Chem. **264:** 20131-20139.

86. **Ito, J., and R. J. Roberts.** 1979. Unusual base sequence arrangements in phage ϕ29 DNA. Gene. **5:** 1-7.

87. **Jacobs, M., H. Koornhof, R. Robins-Browne, C. Stevenson, A. Vermaak, I. Freiman, G. B. Miller, M. A. Witcomb, M. Isaacson, J. I. Ward, and R. Austrian.** 1978. Emergence of multiply resistant pneumococci. N. Engl. J. Med. **299:** 735-740.

88. **Jentsch, S.** 1983. Restriction and modification in *Bacillus subtilis*: sequence specificities of restriction-modification systems *Bsu*M, *Bsu*E, and *Bsu*F. J. Bacteriol. **156:** 800-808.

89. **Jerpseth, B., A. Greener, J. M. Short, and P. L. Kertz.** 1993. New restriction-minus derivatives of XL1-Blue *E. coli* cells. *Strategies*. **6:** 24.

90. **Johnston, R. B., Jr.** 1991. Pathogenesis of pneumococcal pneumonia. Rev. Infect. Dis. **13** (Suppl): S509-S517.

91. **Joklik, W. K., P. Willet, D. B. Amos, C. M. Wilfert.** (eds.). 1988. Streptococcus pneumoniae. pp: 368-377. In " Zinsser Microbiology, nineteenth edition". Appleton and Lange, Norwalk.

92. **Kagan, R. M., and S. Clarke.** 1994. Widespread occurrence of three sequence motifs in diverse S-adenosylmethinine-dependent methyltransferases suggests a common structure for these enzymes. Arch. Biochem. Biophys. **310:** 417-427.

93. **Kahmann, R.** 1984. The *mom* gene of bacteriophage Mu. pp. 29-48. Current topics in Microbiology and Immunology. In "Methylation of DNA", T.A. Trautner.(ed.)., Springer-Verlag, Newyork.

94. **Kahmann, R., and S. Hattman.** Regulation and expression of the *mom* gene. 1987. pp. 93-109. In "Phage Mu", Symonds, N., Toussaint, van de Putte, P., and Howe, M. M., (eds.)., Cod Spring Harbor Laboratory, Cold Spring Harbor, New York.

95. **Karreman, C., and A. de Waard.** 1990. *Agmenellum quadruplicarum* M.*Aqu*I, a novel modification methylase. J. Bacteriol. **172:** 266-272.

96. **Kellerman, O. K., and T. Ferenci.** 1982. Maltose binding protein from *E. coli*. Methods in Enzymology. **90:** 459-463.

97. **Kilic, A. O., M. N. Vijayakumar, and S. F. Al-Khaldi. 1994.** Identification of transfer-related genes in the streptococcal conjugative transpsoson, Tn*5252*. J. Bacteriol. **176:**5145-5150.

98. **Kiss, A., G. Posfai, C. C. Keller, P. Venetianer, and R. J. Roberts.** 1985. Nucleotide sequence of the *Bsu*RI restriction-modification system. Nucleic Acids Res. **13:** 6403-6420.

99. **Kleckner, N., D. Morisato, D. Roberts, J. Bender.** 1984. Mechanism of regulation of Tn*10* transposition. Cold Spring Harb. Symp. Quant. Biol. **49:** 235-244

100. **Klimašauskas, S., A. Timinskas, S. Menkevicius, D. Butkiene, V. Butkus and A. Janulaitis.** 1989. Sequence motifs characteristic of DNA [cytosine-N4] methyltransferases: similarity to adenine and cytosine-C5 methylases. Nucleic Acids Res. **17:** 9823-9832.

101. **Klimašauskas, S., J. L. Nelson, and R. J. Roberts.** 1991. The sequence specificity domain of cytosine-C5 methylases. Nucleic Acids Res. **19:** 6183-6190.

102. **Klimasauskas, S., S. Kumar, R. J. Roberts, and X. Cheng.** 1994. *Hha*I methyltransferase flips its target base out of the DNA helix. Cell. **76:** 357-369.

103. **Knapp, J. S., S. R. Johnson, J. M. Zenilman, M. C. Roberts, and S. A. Morse.** 1988. High-level tetracycline resistance resulting from *tet*M in strains of *Neisseria* spp., *Kingella denitrificans*, and *Eikenella corrodens*. Antimicrob. Agents Chemother. **32:** 765-767.

104. **Kossykh, V. G., A. V. Repyk, and S. Hattman.** 1993. Sequence motifs common to the *Eco*RII restriction endonuclease and the proposed sequence specificity domain of three DNA- (cytosine-C5) methyltransferases. Gene **125:** 65-68.

105. **Kruger, D. H., and T. A. Bickle.** 1983. Bacteriophage survival: multiple mechanisms for avoiding the DNA restriction systems of their hosts. Microbiol. Rev. **47:** 345-360.

106. **Kruger, D. H., C. Schroeder, M. Reuter, I. G. Bogdarina, Y. I. Buryanov, and T. A. Bickle.** 1985. DNA methylation of bacterial viruses T3 and T7 by different DNA methylases in *Escherichia coli* K12 cells. Eur. J. Biochem. **150:** 323-330.

107. **Kruger, D. H., G. J. Barcak, M. Reuter,and H. O. Smith.** 1988. *Eco*RII can be activated by cleavage refractory DNA recognition sites. Nucleic Acids Res. **16:** 3997-4008.

108. **Labahn, J., J. Granzin, G. Schluckebier, D. P. Robinson, W. E. Jack, I. Schildkraut, and W. Saenger.** 1994. Three-dimensional structure of the adenine specific DNA methyltransferase M.*Taq*I in complex with the cofactor S-adenosylmethionine. Proc. Natl. Acad. Sci. USA. **91:**10957-10961.

109. **Lacks, S. A., P. Lopez, B. Greenberg and M. Espinosa.** 1986. Identification and analysis of genes for tetracycline resistance and replication functions in the broad-host-range plasmid pLS1. J. Mol. Biol. **192:**753-765.

110. **Lacks, S. A., and S. S. Springhorn.** 1984. Cloning in *Streptococcus pneumoniae* of the gene for *Dpn*II DNA methylase. J. Bacteriol. **157:** 934-936.

111. **Laemmli, U. K.** 1970. Cleavage of structural proteins during the assembly of the head of the bacteriophage T4. Nature. **227:**680-685.

112. **Lange, C., A. Jugel, J. Walter, M. Noyer-Weidner and T. A. Trautner.** 1991. 'Pseudo' domains in phage-encoded DNA methyltransferases. Nature. 352: 645-648.

113. **Lange, C., M. Noyer-Weidner, T. A. Trautner, M. Weiner and S. A. Zahler.** 1991. M.*H2*I, a multispecific 5C-DNA methyltransferase encoded by *Bacillus amyloliquifaciens* phage H2. Gene. **100:** 213-218.

114. **Laskey, R. A., and A. D. Mills.** 1975. Quantitative film detection of ^{3}H and ^{14}C in polyacrylamide gels by fluorography. Eur. J. Biochem. **56:**335-341.

115. **Lauster, R.** 1989. Evolution of type II DNA methyltransferases. A gene duplication model. J. Mol. Biol. **206:** 313-321.

116. **Lauster, R., T. A. Trautner, and M. Noyer-Weidner.** 1989. Cytosine-specific type II DNA methyltransferases. A conserved enzyme core with variable target-recognizing domains. J. Mol. Biol. **206:** 305-312.

117. **Lautenberger, J. A., and S. Linn.** 1972. The deoxyribonucelic acid modification and restriction enzymes of *Escherichia coli* B. I. Purification, sub-unit structure and catalytic properties of the modification methylase. J. Biol. Chem. **247:** 6176-6182.

118. **Lawrie, J. M., J. S. Downard, and H. R. Whiteley.** 1978. *Bacillus subtilis* bacteriophages SP82, SP01 and ϕe: a comparison of DNAs and peptides synthesized during infection. J. Virol. **27:** 725-737.

119. **Le Bouguenec, C., G. De Cespedes and T. Horaud.** 1988. Molecular analysis of a composite chromosomal conjugative element (Tn*3701*) of *Streptococcus pyogenes*. J. Bacteriol. **170:**3930-3936.

120. **LeBouguenec, C., G .DeCespedes, and T. Horaud.** 1990. Presence of chromosomal elements resembling the composite structure Tn3701 in Streptococci. J. Bacteriol. **172:** 727-734.

121. **LeBouguenec, C., T. Horaud, G. Bieth, R. Colimon, and C. Dauguet.**1984. Translocation of antibiotic resistance markers of a plasmid-free *Streptococcus*

pyogenes (group A) strain into different strepotococcal hemolysin plasmids. Molec. Gen. Genet. **194:** 377- 387.

122. **Lee, K., K. Kam, and P. Shaw.** 1995. A bacterial methyltransferase M.*Eco*HK311 requires two proteins for *in vitro* methylation. Nucleic Acids Res. **23:** 103-108.

123. **Lee, N. S., O. Rutebuka., T. Arakawa., T. A. Bickle., and J. Ryu.** 1997. *Kpn*AI, a new type I restriction-modification system in *Klebsiella pneumoniae*. J. Mol. Biol. **271:** 342-248.

124. **Loenen, W., and N. Murray.** 1986. Modification enhancement by restriction alleviation protein (Ral) of bacteriophage lambda. J. Mol. Biol. **190:** 11-22.

125. **Lopez, P., M. Espinosa, D. L. Stassi, and S. A. Lacks.** 1982. Facilitation of plasmid transfer in *Streptococcus pneumoniae* by chromosomal homology. J. Bacteriol. **150:** 692-701.

126. **Macrina, F. L., R. P. Evans, J. A. Tobian, D. L. Hartley, D. B. Clewell, and K. R. Jones.** 1983. Novel shuttle plasmid vehicles for *Escherichia- Streptococcus* transgeneric cloning. Gene. **25:**145-150.

127. **Makino, O., H. Saito, and T. Ando.** 1980. *Bacillus subtilis*-phage ϕ1 overcomes host-controlled restriction by producing *Bam*Nx inhibitor protein. Mol. Gen. Genet. **179:** 463-468.

128. **Makino, O., J. Kawamura, H. Saito, and Y. Ikada.** 1979. Inactivation of restriction endonuclease BamNx after infection with phage ϕNR2. Nature. **277:** 64-66.

129. **Malone, T., R. M. Blumenthal, and X. Cheng.** 1995. Structure-guided analysis reveals nine sequence motifs conserved among DNA amino-methyltransferases, and suggests a catalytic mechanism for these enzymes. J. Mol. Biol. **253:** 618-632.

130. **Mannarelli, B. M., Balganesh, T. S., Greenberg, B., Springhorn, S. S., and Lacks, S. A.** 1985. Nucleotide sequence of the *Dpn*II DNA methylase gene of *Streptococcus pneumoniae* and its relationship to the *dam* gene of *Escheichia coli*. Proc. Natl. Acad. Sci. **82:** 4468-447.

131. **Marinus, M.G.** 1987. DNA Methylation in *Escherichia coli*. Ann. Rev. Genet. 1987. **21:** 113-131.

132. **Marmur, J.** 1961. A procedure for the isolation of deoxyribonucleic acid from microorganisms. J. Mol. Biol. **3:**208-21

133. **McClelland, M.**1985. Selection against *dam* methylation sites in the genomes of DNA of enterobacteriphages. J. Mol. Evol. **21:** 317-322.

134. **McCommas, S. A., and Syvanen, M.** 1988. Temporal control of transposition in Tn*5*. J. Bacteriol. **170:** 889-894.

135. **Meisel, A., T. A. Bickle, D. H. Kruger, and C. Schroeder.** 1992. Type III restriction enzymes need two inversely oriented recognition sites for DNA cleavage. Nature. **355:** 467-469.

136. **Meister, J., M. MacWilliams, P. Hubner, H. Jutte, E. Skrzypek, A. Piekarowicz, and T. A. Bickle.** 1993. Macroevolution by transposition: drastic modification of DNA recognition by a type I restriction enzyme following Tn*5* transposition. EMBO J. **12:** 4585- 4591.

137. **Mi, S., and R. J. Roberts.** 1993. The DNA binding affinity of *Hha*I methylase is increases by a single amino acid substitution in the catalytic center. Nucleic Acids Res. **21:** 2459-2464.

138. **Miyamura, S., H. Ochiai, Y. Nitahara, Y. Nakagawa, and M. Terao.** 1977. Resistance mechanism of chloramphenicol in *Streptococcus hemolyticus*, *Streptococcus pneumoniae*, and *Streptococcus faecalis*. Microbiol. Immunol. **21:** 69-76.

139. **Moffatt, B. A., and F. W. Studier.** 1987. T7 lysozyme inhibits transcription by T7 RNA polymerase. Cell. **49:** 221

140. **Murray, N. E., J. A. Gough., B. Suri., and T. A. Bickle.** 1982. Structural homologies among type-I restriction-modification systems. EMBO J. **1:** 535-539.

141. **Musher, D. M.**1992. Infections caused by *Streptococcus pneumoniae*: clinical spectrum, pathogenesis, immunity, and treatment. Clin. Infect. Dis, **14:** 801-809.

142. **Najar, U. M., and M. N. Vijayakumar.** In preparation.

143. **Noyer-Weidner, M., and T. A. Trautner.**: Methylation of DNA in prokaryotes. 1993. pp. 39-108. In "DNA methylation: Molecular biology and biological significance", Jost, J. P., and Saluz, H. P. (eds.). Birkhhäuser Verlag Basel, Switzerland .

144. **O'Gara, M., K. McCloy, T. Malone, and X. Cheng.** 1995. Structure-based sequence alignment of three AdoMet-dependent methyltransferases. Gene. **157:** 135-138.

145. **Piekarowicz, A.** 1982. *Hine*I is an isoschizomer of *Hinf*III restriction endonuclease. J. Mol. Biol. **157:** 373-381.

146. **Plasterk, R. H. A., Vrieling, H., and Van de Putte, P.** 1983. Transcription initiation of Mu *mom* depends on methylation of the promoter region and a phage-coded transactivator. Nature. **301:** 344-347.

147. **Posfai, J., A. S. Bhagwat, G. Posfai, and R. J. Roberts.** 1989. Predictive motifs derived from methyltransferases. Nucleic Acids Res. **17:** 2421-2435.

148. **Posfai, G., S. C. Kim, L. Szilak, A. Kovacs, and P. Venetianer.** 1991. Complementation by detached parts of GGCC-specific methyltransferases. Nucleic Acids Res. **19:** 4843-4847.

149. **Poyart-Salmeron, C., P. Trieu-Cuot, C. Carlier, and P. Courvalin.** 1989. Molecular characterization of two proteins involved in the excision of the conjugative transposon Tn*1545*: homologies with other site-specific recombinases. EMBO. J. **8:**2425-2433.

150. **Poyart-Salmeron, C., P. Trieu-Cuot, C. Carlier, and P. Courvalin.** 1990. The integration-excision system of the conjugative transposon Tn*1545* is structurally and functionally related to those of lambdoid phages. Mol. Microbiol. **4:**1513-1521.

151. **Raleigh, E. A., and N. Kleckner.** 1986. Quantitation of insertion sequence IS*10* transposase gene expression by a method generally applicable to any rarely expressed gene. Proc. Natl. Acad. Sci. USA. **83:** 1787-1791.

152. **Rauch, P. J. G., and W. M. DeVos.** 1992. Characterization of the novel Nisin-Sucrose Conjugative transposon Tn*5276* and its insertion in *Lactococcus lactis*. J. Bacteriol. **174:**1280-1287.

153. **Rauch, P. J. G., and W. M. de Vos.** 1994. Identification and characterization of genes involved in excision of the *Lactococcus lactis* conjugative transposon Tn*5276*. J. Bacteriol. **176:**2165-2171.

154. **Read, T. D., A. T. Thomas, and B. M. Wilkins.** 1992. Evasion of type I and type II DNA restriction systems by IncI1 plasmid ColIb-P9 during transfer by bacterial conjugation. Mol. Microbiol. **6:** 1933-1941.

155. **Reinisch, K. M., L. Chen, G. L. Verdine, and W. N. Lipscomb.** 1995. The crystal structure of *Hae*III methyltransferase covalently complexed to DNA: an extrahelical cytosine and rearranged base pairing. Cell. **82:** 143-153.

156. **Rice, L. B., and L. L. Carias.** 1994. Studies on excision of conjugative transposons in *Enterococci*: evidence for joint sequences composed of strands with unequal numbers of nucleotides. Plasmid. **31:**312-316.

157. **Roberts, D., B. C. Hoopes, W. R. McClure, and N. Kleckner.** 1985. IS*10* transposition is regulated by adenine methylation. Cell **43:** 117-130.

158. **Roberts, M.C.**1990. Chracterization of the TetM determinants in urogenital and respiratory bacteria. Antimicrob. Agents. Chemother. **32:** 476-478.

159. **Robins-Browne, R., Gaspar, M. N., Ward, J. I., Wachsmuth, I. K., Koornhof, H. J., Jacobs, M. R., and Thornsberry, C.** 1979. Resistance mechanisms of multiply resistanct pneumococci: antibiotic degradation studies. Antimicrob. Agents Chemother. **15:** 470- 474.

160. **Sabelnikov, A. G., B. Greenberg, and S. A. Lacks.** 1995. An extended -10 promoter alone directs transcription of the *Dpn*II operon of *Streptococcus pneumoniae*. J. Mol. Biol. **250:** 144-155.

161. **Salyers, A. A., N. B. Shoemaker, A. M. Stevens, and L. Li.** 1995. Conjugative transposons: an unusual and diverse set of integrated gene transfer elements. Microbiol. Rev. **59:** 579-590.

162. **Salyers, A. A., N. B. Shoemaker, and E. P. Guthrie.** 1987. Recent advances in *Bacteroides* genetics. Crit. Rev. Microbiol. **14:** 49-71.

163. **Sambrooks, J., E. F. Fritsch, and N. T. Maniatis.** (ed.). 1989. Molecular Cloning. A Laboratory Manual. 2nd ed. Cold Spring Harbor Laboratory Press.

164. **Sampath, J., and M. N. Vijayakumar.** 1998. Identification of a DNA cytosine methyltransferase gene in Conjugative transposon Tn*5252*. Plasmid. **39:** 63-76.

165. **Santi, D. V., and L. W. Hardy.** 1987. Catalytic mechnism and inhibition of tRNA(Uracil-5) methyltransferase: evidence for covalent catalysis. Biochemistry. 26: 8599-8606.

166. **Santi, D. V., and P. V. Danenberg.** 1984. Folates in pyrimidine nucleotide biosynthesis. pp. 343-396. In "Folates and Pterins, Volume 1", Blakely R. L., and S. J. Benkovic (eds.). Wiley- Interscience, New York.

167. **Saunders, C. W., and W. R. Guild.** 1980. Properties and Transforming Activities of two plasmids in *Streptococcus pneumoniae*. Mol. Gen. Genet. **180:**573-578.

168. **Schluckbier, G., M. O'Gara, W. Saenger, and X. Cheng.** 1995. Universal catalytic domain structure of AdoMet-dependent methyltransferases. J. Mol. Biol. **247:** 16-20.

169. **Schouler, C., F. Cleir, A. L. Lerayer, S. D. Erlich, and M. C. Chopin.**1998. A type Ic restriction-modification system in *L.lactis*. J. Bacteriol. **180:** 407-411.

170. **Schroeder, C., H. Jurkschat, , A. Meisel, J. G. Reich, and D. Kruger.** 1986. Unusual occurrence of EcoP1 and EcoP15 recognition sites and counterselection of type II methylation and restriction sequences in bacteriophage T7 DNA. Gene. **45:** 77-86.

171. **Scott, J. R.** 1992. Sex and the single circle: conjugative transposition. J. Bacteriol. **174:** 6005-6010.

172. **Scott, J. R.** 1993. Conjugative transposons. pp.597-614. In "*Bacillus subtilis* and other Gram-positive bacteria", Sonenshein, A. L., Hoch, J. A., and Losick, R. (eds.). American Society for Microbiology, Washington, D. C .

173. **Scott, J. R., P. A. Kirchman, and M. G. Caparon.** 1988. An intermediate in transposition of the conjugative transposon Tn*916*. Proc. Natl. Acad. Sci. USA. **85:** 4809-4813.

174. **Sharp, P. M. 1986.** Molecular evolution of bacteriophages: evidence of selection against recognition sites of host restriction enzymes. Mol. Biol. Evol. **3:** 75-83.

175. **Shoemaker, N. B., and W. R. Guild.** 1974. Destruction of low efficiency markers is a slow process occuring at a heteroduplex stage of transformation. J. Bacteriol. **128:** 283-290.

176. **Shoemaker, N. B., M. D. Smith, and W. R. Guild.** 1979. Organization and transfer of heterologous chloramphenicol and tetracycline resistance genes in pneumococcus. J. Bacteriol. **139:** 432-441.

177. **Shoemaker, N. B., M. D. Smith, and W. R. Guild.** 1980. DNase-resistant transfer of chromosomal *cat* and *tet* insertions by filter mating in pneumococcus. Plasmid. **3:**80-87.

178. **Simons, R. W., and N. Kleckner.** 1983. Translational control of IS*10* transposition. Cell. **34:** 683-691.

179. **Simons, R. W., B. C. Hoopes, W. R. McClure, and N. Kleckner.** 1983. Three promoters near the termini of IS*10*: pIN, pOUT, and PIII. Cell. **34:** 673-682.

180. **Sitaram, R., and K. Dybvig.** 1997. The *hsd* loci of *Mycolplasma pulmonis*: organization, rearrangement and expression of gene. Mol. Microbiol. **26:** 109-120.

181. **Smith, M. D.** 1981. The Genetics of mutltiple drug resistance in *Streptococcus pneumoniae*. Ph. D. Thesis. Department of Biochemistry. Duke University.

182. **Smith, M. D., and W. R. Guild.** 1979. A plasmid in *Streptococcus pneumoniae*. J. Bacteriol. **137:** 735-739.

183. **Smith, M. D., and W. R. Guild.** 1980. Improved method for conjugative transfer by filter mating in *Streptococcus pneumoniae*. J. Bacteriol. **144:** 457-457.

184. **Smith, M. D., and W. R. Guild.** 1982. Evidence for transposition of the conjugative R determinants of *Streptococcus agalactiae* B109. pp. 109-111. In "Microbiology-1982", D. Schlessinger (ed.). American Society for Microbiology, Washington, D.C.

185. **Smith, M. D., N. B. Shoemaker, V. Burdett, and W. R. Guild.** 1980. Transfer of plasmids by conjugation in *Streptococcus pneumoniae*. Plasmid. **3:** 70-79.

186. **Southern, E.** 1975. Detection of specific sequences among DNA fragments separated by gel electrophoresis. J. Mol. Biol. **98:** 503-517.

187. **Spoerel, N., P. Herrlich, and T. A. Bickle.** 1979. A novel bacteriophage defence mechanism: the anti-restriction protein. Nature. **278:** 30-34.

188. **Srinivas, P.** 1997. Structural and Functional analysis of a regulatory region in the streptococcal conjugative transposon, Tn*5252*. Ph.D. Thesis. Oklahoma State University.

189. **Srinivas, P., A. O. Kilic, and M. N. Vijayakumar.** 1997. Site-specific nicking *in vitro* at *ori*T by the DNA relaxase of Tn*5252*. Plasmid **37:** 42-50.

190. **Stassi, D. L., P. Lopez, M. Espinosa, and S. A. Lacks.** 1981. Cloning of Chromosomal Genes in *Streptococcus pneumoniae*. Proc. Nat. Acad. Sci., U.S.A. **78:** 7028-7032.

191. **Stephenson, F. H., B. T. Ballard, H. W. Boyer, J. M. Rosenberg, and P. J. Greene.** 1989. Comparison of the nucleotide and amino acid sequences of the *Rsr*I and *Eco*RI restriction endonucleases. Gene **85:** 1-13.

192. **Sternberg, N.** 1985. Evidence that adenine methylation influences DNA-Protein interactions in *E. coli*. J. Bacteriol. **164(1):** 490-493.

193. **Sternberg, N., and J. Coulby.** 1990. Cleavage of the bacteriophage P1 packaging site (*pac*) is regulated by adenine methylation. Proc. Natl. Acad. Sci. USA. **87:** 8070-8704.

194. **Sternberg, N., B. Sauer, R. Hoess, and K. Abermski.** 1986. Bacteriophage P1 *cre* gene and its regulatory region. J. Mol. Biol. **187:** 197-212.

195. **Studier, F. W.** 1991. Use of bacteriophage T7 lysozyme to improve an inducible T7 expression system. J. Mol. Biol. **219:** 37- 44.

196. **Studier, F. W., and N. R. Movva.** 1976. SAMase gene of bacteriophage T3 is responsible for overcoming host restriction. J. Virol. **19:** 136-145.

197. **Su, Y. A., and D. B. Clewell.** 1993. Characterization of the left four kilobases of conjugative transposon, Tn*916*. Determinants involved in excision. Plasmid. **30:** 234-250.

198. **Sullivan, K. M., and J. R. Saunders.** 1989. Nucelotide sequence and genetic organization of the *Ngo*PII restriction-modification system of *Neisseria gonorrhoeae*. Mol. Gen. Genet. **216:** 380-387.

199. **Suri, B., V. Nagaraja., T. A. Bickle.** 1984. Bacterial DNA methylation. pp. 1-10. In "Methylation of DNA", Trautner, T. A. (ed.). Current topics in Microbiology and Immunology vol. 108. Springer-Verlag, New York.

200. **Swartley, J. S., C. F. McAllister, R. A. Hajjeh, D. W. Heinrich, and D. S. Stephens.** 1993. Deletions of Tn*916*-like transposons are implicated in *tet*M-mediated resistance in pathogenic *Neisseria*. Mol. Microbiol. **10:** 299-310.

201. **Titheradge, A. T., D. Ternent., N. E. Murray.** 1996. A third family of allelic *hsd* genes in *Salmonella enterica*: sequence comparisons with related protein to identify conserved regions implicated in restriction of DNA. Mol. Microbiol. **22:** 437-447.

202. **Tran-Betcke, A., B. Behrens, M. Noyer-Weidner, and T. A. Trautner.** 1986. DNA methyltransferase genes of *Bacillus subtilis* phages: comparison of their nucleotide sequences. Gene **42:** 89-96.

203. **Trautner, T.A., and M. Noyer-Weidner.** 1993. Restriction-modification and methylation systems in *Bacillus subtilis*, related species, and their phages. pp. 539-552. In "*Bacillus subtilis* and other Gram positive bacteria, biochemistry, physiology and molecular genetics". Sonenshein, A. C., J. A. Hoch, and R. Losick. (eds.). American Society for microbiology, Washington, D. C.

204. **Trautner, T. A., T. S. Balganesh, and B. Pawlek.** 1988. Chimeric multispecific DNA methyltransferases with novel combination of target recognition. Nucleic Acids Res. **16:** 6649- 6657.

205. **Trieu-Cuot, P., and P. Courvalin.** 1983. Nicleotide sequence of the *Streptococcus faecalis* plsmid gene encoding the 3'5''-aminoglycoside phosphotransferase type III. Gene. **23:** 331-341.

206. **Vijayakumar, M. N., and S. Ayalew.** 1992. Nucleotide sequence analysis of the termini and chromosomal loci involved in the site-specific integration of the streptococcal conjugative transposon, Tn*5252*. J. Bacteriol. **175:**2713-2719.

207. **Vijayakumar, M. N., S. D. Priebe, and W. R. Guild.** 1986. Structure of a conjugative element in *Streptococcus pneumoniae*. J. Bacteriol. **166:** 978-984.

208. **Vijayakumar, M. N., S. D. Priebe, G. Pozzi, J. M. Hageman, and W. R. Guild.** 1986. Cloning and physical characterization of chromosomal conjugative elements in streptococci. J. Bacteriol. **166:** 972-977.

209. **Vovis, G. F., K. Horiuchi, and N. D. Zinder.** 1974. Kinetics of methylation by a restriction endunuclease from *Eschrechia coli* B. Proc. Natl. Acad. Sci. USA. **71:** 3810-3813.

210. **Waldor, M. K., H. Tschape, and J. J. Mekalanos.** 1996. A new type of conjugative transposon encodes resistance to Sulfamethoxazole, Trimethoprim and Streptomycin in *Vibrio cholerae* 0139. J. Bacteriol. **178:** 4157-4165.

211. **Walter, J., M. Noyer-Weidner, and T. A. Trautner.** 1990. The amino acid sequence of the CCGG recognizing DNA methyltransferase M.*Bsu*FI: implications

for the analysis of sequence recognition by cytosine DNA methyltransferases. EMBO J. **9:** 1007-1013.

212. **Walter, J., T. A. Trautner, and M. Noyer-Weidner.** 1992. High plasticity of multispecific DNA methyltransferases in the region carrying DNA target recognizing enzyme modules. EMBO J. **11:** 4445-4450.

213. **Weisblum, B.** 1974. In "Microbiology 1974", Schlessinger, D. (ed.). American Society for Microbioliology, Washington, D. C.

214. **Wilke, K., E. Pauhut, M. Noyer-Weidner, R. lauster, B. Pawlek, B. Behrens, and T. A. Trautner.** 1988. Sequential order of target-recognizing domains in multispecific DNA-methyltransferases. EMBO J. **7:** 2601-2609.

215. **Wilson, G. G.** 1988. Cloned restriction-modification systems- a review. Gene. **74:** 281-289.

216. **Wilson, G. G.** 1992. Amino acid sequence arrangements of DNA-methyltransferases. Methods Enzymol. **216:** 259-279.

217. **Wison, G. G., and N. E. Muray.** 1991. Restriction and modification systems. Annu. Rev. Genet. **25:** 585-627.

218. **Wohlhieter, J. A., J. R. Lazere, N. J. Snellings, E. M. Johnson, R. M. Synenki, and L. S. Baron.** 1975. Characterization of transmissible genetic elements from sucrose-fermenting *Salmonella* strains. J. Bacteriol. **122:** 401-406.

219. **Wu, J. C., and D. V. Santi.** 1985. On the mechanism and inhibition of DNA cytosine methyltransferases. . pp. 119-129. In "Biochemistry and Biology of DNA Methylation", G. I Cantoni and A. Razin (eds.). Alan R. Liss, Inc. New York.

220. **Wyszynski, M.W., S. Gabbara, A. S. Bhagwat.** 1992. Substitutions of a cysteine conserved among DNA cytosine methylases result in a variety of phenotypes. Nucleic Acids Res. **20:** 319-326.

221. **Xu, G. –L., W. Kapfer, J. Walter, and T. A. Trautner.** 1992. *Bsu*BI –an isotypic restriction and modification system of *Pst*I: characterization of the *Bsu*BI genes and enzymes. Nucleic Acids Res. **20:** 6517-6523.

222. **Yin, J. C. P., M. P. Krebs, and W. S. Reznikoff.** 1988. Effect of *dam* methylation on Tn*5* transposition. J. Mol. Biol. **199:** 35-45.

223. **Zighelboim, S., and A. Tomasz.** 1980. Penicillin-binding proteins of multipley abtibiotic-resistant South African strains of *Streptococcus pneumoniae*. Antimicrob. Agents Chemother. **17:** 434-442.

VITA

JANARDHAN SAMPATH

Doctor of Philosophy

Thesis: IDENTIFICATION AND CHARACTERIZATION OF A DNA 5^{m}C-CYTOSINE METHYLTRANSFERASE GENE IN STREPTOCOCCAL CONJUGATIVE TRANSPOSON Tn*5252*.

Major Field: Microbiology.

Biographical:

Personal Data: Born in Srirangam, Trichy, India, July 27, 1968, son of N. Sampath and Saroja.

Education: Graduated from E.R.Higher Secondary School, Trichy, India in June 1985; received Bachelor of Pharmacy and Master of Biological Sciences Degrees from Birla Institute of Technology and Science, in December 1990; completed requirements for the Doctor of PhilosophyDegree at Oklahoma State University in December 1998.

Professional Experience: Medical representative, Reckitt and Coleman India Ltd, May – August 1991. Quality Assurance Trainee, Nicholas India Ltd, August 1991 to December 1992. Research and Teaching Assistant, Department of Microbiology and Molecular Genetics, Oklahoma State University, January 1993 to December 1998.

Publications:

Ray, S. K., S. Rudhra., R. Chandrasekar., K.L. Premkanth., M.V.S. Murthy., **J. Sampath.,** S. Venkatesan. 1991. Design of potent anti-fungal cinnamic acid derivatives. A comparative multiparametric QSAR study involving newly developed electronic parameters, Residual Information Flow and Residual Electronic Force. Medical Science Research. 19: 759-763.

Alarcon-Chaidez, F. J., **J. Sampath.,** P. Srinivas., and M. N. Vijayakumar. 1996. Tn*5252*: A model for complex Streptococcal Conjugative transposons. p. 1029-1032. In "*Streptococci* and the host". Horaud, T., A. Bouvet, R. Leclercq, H. DeMontclos, and M. Sicard (eds.). Plenum Publishing, New York.

Sampath, J., and M. N. Vijayakumar. 1998. Identification of a DNA cytosine methyltransferase gene in Conjugative transposon Tn*5252*. Plasmid. **39:** 63-76.

Professional memberships: American Society for Microbiology, Sigma Xi.